DOCUMENT RESUME

ED 252 810 CS 006 420

AUTHOR	Kamil, Michael L., Ed.; Boswick, Mary M., Ed.
TITLE	Directions in Reading: Research and Instruction. Thirtieth Yearbook of the National Reading Conference.
INSTITUTION	National Reading Conference, Inc.
PUB DATE	81
NOTE	306p.; Proceedings of the Annual Meeting of the National Reading Conference (30th, San Diego, CA, December 3-6, 1980).
PUB TYPE	Books (010) -- Collected Works - Conference Proceedings (021) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC13 Plus Postage.
DESCRIPTORS	Adult Literacy; Adults; Beginning Reading; Cognitive Processes; Early Childhood Education; Early Reading; Elementary Secondary Education; Instructional Materials; Language Processing; Reading Comprehension; *Reading Instruction; *Reading Processes; *Reading Research; Reading Strategies; Reading Tests; Teaching Methods
IDENTIFIERS	Orthography; Text Structure; Theory Practice Relationship

ABSTRACT

Reflecting the increasing complexity and sophistication in reading research, the papers in this collection represent a wide spectrum of approaches, philosophies, viewpoints, and techniques of scholarly endeavor in their treatment of reading research and instruction. Following an address, by the president of the organization, on reading comprehension research in search of classroom validation, the 36 articles presented at the 1980 meeting of the National Reading Conference are arranged according to the following categories: (1) strategies in reading, (2) orthography and reading, (3) theoretical concerns, (4) instructional concerns, (5) comprehension and test structure, (6) adjunct aides, (7) early reading, and (8) adult reading. The volume concludes with the program from the meeting. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED252810

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

● Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

DIRECTIONS IN READING: RESEARCH AND INSTRUCTION

Thirtieth Yearbook
of
The National Reading Conference

Michael L. Kamil, Editor
University of Illinois at Chicago Circle

with the editorial
assistance of
Mary M. Boswick
University of Illinois at Chicago Circle

Published by
The National Reading Conference, Inc.

1981

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
NRC

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

CS 006 420

Copyright, 1981
THE NATIONAL READING CONFERENCE, INC.
All rights reserved. This book or parts thereof
may not be reproduced in any form without
permission of the publishers

Distributed by
THE NATIONAL READING CONFERENCE
1230 17th Street, N.W.
Washington, DC 20036

The ideas expressed in the papers published
in the yearbook are those of the authors and do
not necessarily represent the viewpoints
of the National Reading Conference

Information concerning membership in The National Reading Conference a subscription to *Journal of Reading Behavior* and the *NRC Yearbook* may obtained by writing to:

National Reading Conference
1230 17th Street, N.W.
Washington, DC 20036

OFFICERS AND BOARD OF DIRECTORS
OF THE
NATIONAL READING CONFERENCE, INC.

FRANK GREENE, *President*
McGill University

IRENE ATHEY, *Vice-President*
Rutgers University

M. TRIKA SMITH-BURKE, *Treasurer*
New York University

JANET J. LARSEN, *Secretary*
University of Florida

HARRY SINGER, *Past President*
University of California-Riverside

THOMAS ESTES, *Board Member*
University of Virginia

JOHN T. GUTHRIE, *Board Member*
International Reading Association

THOMAS G. STICHT, *Board Member*
Human Resources Research Organization

JOANNA P. WILLIAMS, *Board Member*
Columbia University

MICHAEL McKENNA, *Convention Manager*
Wichita State University

MICHAEL L. KAMIL, *Yearbook Editor*
University of Illinois at Chicago Circle

FRED W. OHNMACHT AND RICHARD ALLINGTON, Editors,
Journal of Reading Behavior
State University of New York at Albany

PRISCILLA DRUM, *Publications Chairperson*
University of California at Santa Barbara

EDITORIAL ADVISORY BOARD

Patricia L. Anders
University of Arizona

Emery P. Bliesmer
Pennsylvania State University

Connie Bridge
University of Kentucky

Robert C. Calfee
Stanford University

Eugene H. Cramer
University of Illinois at
Chicago Circle

Carol Davis
Purdue University

Priscilla Drum
University of California at
Santa Barbara

Gerald G. Duffy
Michigan State University

Edward Fry
Rutgers University

Jane Hansen
University of New Hampshire

Joyce E. Hood
University of Iowa

Dale D. Johnson
University of Wisconsin at
Madison

Judith A. Langer
University of California at
Berkeley

Ron C. Leslie
New York University

Terry Lovelace
University of Southwestern
Louisiana

Mary Beth Marr
State University of New York at Albany

Alden J. Moe
Purdue University

Allan Nielsen
University of Calgary

Edward Paradis
University of Wyoming

P. David Pearson
University of Illinois at
Champaign

Charles W. Peters
Oakland Public Schools

Susanna Pflaum
University of Illinois at
Chicago Circle

Lenore Ringler
New York University

Diane Schallert
University of Texas

Timothy Shanahan
University of Illinois at
Chicago Circle

Harry Singer
University of California at
Riverside

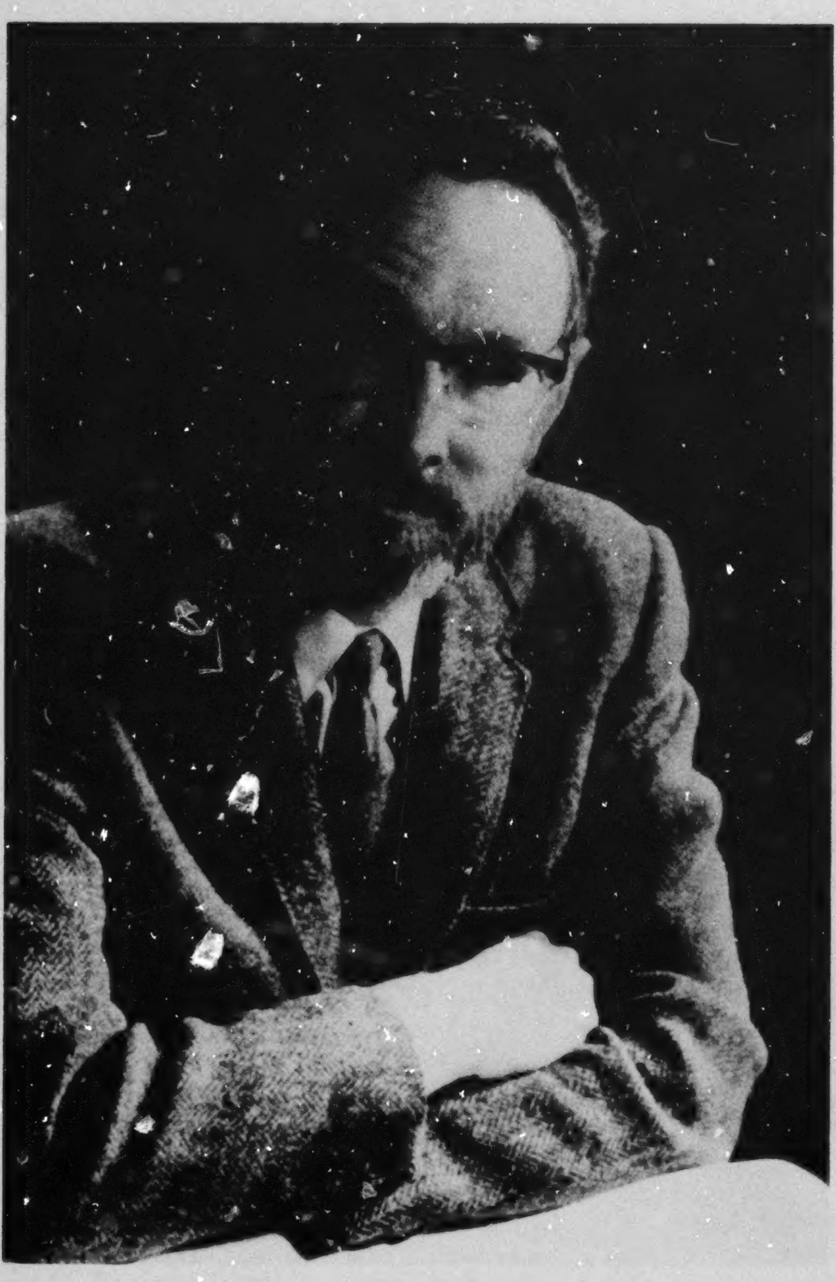
Dixie Lee Spiegel
University of North Carolina

Elizabeth Sulzby
Northwestern University

Robert J. Tierney
University of Illinois at
Champaign

John Towner
Western Washington University

Robert A. Wisher
Office of Naval Research



EDWARD B. FRY

Oscar Causey Award

The Oscar S. Causey Award is made only to those persons who have made distinguished contributions to the field of reading and have given exemplary and dedicated service to the National Reading Conference.

Dr. Edward B. Fry is a most worthy recipient of this award. He is the author of several highly acclaimed books which have been translated into Spanish, Italian, Dutch and Czech. His text *Reading for Classroom and Clinic* (McGraw Hill, 1972) is widely used in graduate reading courses around the nation, while an earlier book, *Teaching Faster Reading* (Cambridge University Press, 1963) is used in reading improvement courses in Great Britain, Africa, and Latin America.

Dr. Fry has published a nationally famous graph for determining readability which has been reprinted in a number of textbooks and articles. Additionally, he is the author of over 60 articles on such varied topics as readability, teaching machines, ITA, and preparation of reading teachers. Curriculum materials that he has developed include a typing course for children, filmstrips, overhead transparencies and card reader programs on phonics, basic vocabulary, reading improvement drill books, criterion-referenced tests, and a videotape reading improvement program by Time/Life used in industry and universities. His 1980 *Instant Word List* has attracted wide attention.

Dr. Fry served as President of the National Reading Conference from 1974 through 1976, and on the Board of Directors as Past President during the next two years. In spite of his many other professional activities such as project director for the U.S. Office of Education First Grade Studies, his membership on national committees of APA, AERA, and NEA and on the Board of Directors of the New Jersey Reading Teachers' Association, his interest in NRC has been sustained over more than a decade.

Since 1963, Dr. Fry has been director of the Reading Center at Rutgers University. His academic qualifications, professional activity, and devoted service to NRC are outstanding. The Oscar Causey Award is presented in recognition of his distinguished career in reading.

TABLE OF CONTENTS

PRESIDENTIAL ADDRESS

Hypotheses on reading comprehension research in search of classroom validation 1
Harry Singer

STRATEGIES IN READING

Metacognition: The development of selective attention strategies for learning from texts 21
Ann Brown
The influence of topic familiarity and the author-reader relationship on detection of inconsistent information 44
Taffy Raphael and Robert Tierney
Eye movements reveal components of flexible reading strategies ... 51
Wayne Shebilske & Dennis Fisher

ORTHOGRAPHY AND READING

Letter-sound regularity and orthographic structure 57
Richard Venezky
The role of orthographic redundancy, versatility, and spelling-sound correspondences in word identification 74
Connie Juel & Robert Solso

THEORETICAL CONCERNS

Developmental considerations of reading research: Applications to good and poor reader research 83
Mark Aulls
A comment on Kenneth Goodman 92
Phillip Gough
Questions from a metatheoretical perspective: The interdependence of solutions to issues involved in the development of reading comprehension models 96
Donald J. Leu, Jr.
A partial reading model utilizing language unit size by frequency .. 103
Edward Fry
Methodological problems in researching strategy deficiencies among readers: Some recommendations 108
Ellen Bouchard Ryan

INSTRUCTIONAL CONCERNS

Teacher effectiveness research: Implications for the reading profession 113
Gerald G. Duffy
Diagnostic and prescriptive decisions made by specialists and classroom teachers: Quadratic assignment as a reliability measure 137
Linda M. McHugh & Ray R. Buss
Toward reassessment of the principles underlying the choice of vocabulary and the teaching of word analysis skills in reading instructional material 144
Sandra Stotsky

A comparison of inservice and preservice teacher verbal feedback to student miscues across two difficulty levels of text.......................... 150
James V. Hoffman, Sharon O'Neal & Christopher Baker

COMPREHENSION & TEXT STRUCTURE

The effects of text type and familiarity on the nature of information recalled by readers.. 157
Sandra Berkowitz & Barbara Taylor
Good and poor readers' ability to distinguish among and recall ideas on different levels of importance.. 162
Lea M. McGee
The effects of typographical cueing on memory for superordinate structures in connected discourse... 170
Barbara A. Rennie, Allan R. Neilsen & Carl Braun
Remembering information from text: a comparison of the text structure of educable mentally retarded and average students' recalls......... 175
Candace S. Bos
The effects of induced mental imagery upon comprehension: A comparison of written versus oral presentation................................ 180
Linda B. Gambrell, Patricia S. Koskinen & Jack N. Cole
The effects of textual cohesion on cloze comprehension scores..... 187
Connie A. Bridge, Susan Thomas, Linda Evers
Readers' reactions to temporal disruption in stories........................ 193
Jill Fitzgerald Whaley
Selected prose characteristics of an ESL narrative and their effect on comprehension.. 198
Thomas W. Bean & Ashley Bishop
Anglo and Chicano comprehension of ethnic stories........................ 205
Nancy Rogers-Zegarra & Harry Singer
An investigation of the relationship of cohesion to linguistic marking, discourse structure and content familiarity... 211
Allan R. Neilsen
The effects of linguistic connectives and prior knowledge on comprehension of good and poor readers.. 217
Gregory G. McKenzie, Allan R. Neilsen & Carl Braun

ADJUNCT AIDS

The effects of three types of written introductions on children's processing of text... 221
Tess O. Kosoff
The effect of manipulating the advance organizer and other prereading strategies on comprehension of abstract text................ 230
Frank Spencer, Michael Johnston & Wilbur Ames
The effects of modified skimming on college students' recall and recognition of expository text.. 234
Tina Jacobowitz
The quality, quantity, and variety of teacher and pupil questions and pupil responses in discussions following group mapping activities. 240
Bonnie C. Greenslade

The effects of postreading questions on children's comprehension and learning 245
Karen K. Wixson

EARLY READING

Cognitive development, metalinguistic ability and invented spellings: Comparisons and correlations 251
Jerry Zutell
How preschoolers interact with written communication 259
William H. Teale, Elette Estrada & Alonzo B. Anderson
Three first graders concept of word and concepts about the language of literacy instruction 268
Tobie S. Sanders
Children's concept of temporal sequence and their comprehension of inverse sentences 275
Judith Greenwald

ADULT READING

A critical review of developments in adult literacy 281
Rosemarie J. Park

FOREWORD

Reading research has become increasingly complex and sophisticated. This trend is evident in the papers selected for this volume. It was evident in the difficulty of choosing among the submitted papers.

With complexity and sophistication has come a reaffirmation of the need to translate research into practice. This too is evident in papers that generate hypotheses for classroom practice, synthesize wide ranges of research results, or report the results of applications. All of these are represented in this collection.

This volume would not have been possible without significant contributions from a large number of people. Mary Boswick has been a diligent and reliable assistant. She kept the process organized and goal-directed. The Editorial Advisory Board as always, played an important role by evaluating manuscripts. Martha Taylor helped by guiding the composition and printing. The authors of the papers also helped by being prompt and responsive to need, throughout. The staff at AERA headquarters also provided helped at critical points in the preparation of the *Yearbook*. Finally, my assistants from earlier years deserve much credit and thanks. It was essentially their system that was used for this volume.

Three years of editing the *Yearbook* have been exciting, informative, rewarding, frustrating, and fun. It is with mixed emotions that I pass the job to new editors. I wish them luck and the same kind of help I received from so many.

Michael L. Kamil
Editor, *NRC Yearbook*

HARRY SINGER
University of California,
Riverside

HYPOTHESIS ON RESEARCH COMPREHENSION IN SEARCH OF CLASSROOM VALIDATION

Presidential Address presented at The Annual Meeting of The National Reading Conference, San Diego, California, December 1980.

In the past two decades we have witnessed a virtual explosion of theories and research in the psychology of reading. This explosion came after the topic of reading had almost disappeared from psychological investigations. This up-and-down status of reading reflects the revolutionary changes that have occurred in the field of psychology over the past 100 years. Some landmark research conducted over the last 100 years, shown in Table 1, depicts these changes.

TABLE 1
Psychological Research on Reading
From 1879 to 1980

1879	Javal	Eye movement behavior during reading consists of saccades and return sweep movements.
1886	Cattell	Adults perceive a word more rapidly than its individual letters.
1908	Huey	*Psychology and Pedagogy of Reading*
1917	Thorndike	Analyses of responses in reading leads to definition that reading is reasoning.
1922	Buswell	Eye movement behavior in reading changes over grades 1-12 by becoming more rhythmical, shorter in pause duration, and longer in span of perception.
1922	Judd & Buswell	During silent reading, readers vary their eye-movement behavior according to the difficulty of the material and their purposes in reading.
1932	Bartlett	*Remembering:* Successive recall of Indian story, "War of the Ghosts," demonstrates that reading is a constructive process and recall is a reconstructive process based on interaction between the text data and the readers' schemata and beliefs.
1946	Thurstone	High school reading tests consist of two factors: Word recognition and reasoning.
1952	Anderson & Dearborn	*Psychology of Reading.* Learning to read consists of formation of S-R associations and then using these associations to perceive printed words.
1953	Holmes	Reading consists of two interrelated components: speed and power. Underlying each component is an interrelated hierarchy of substrata factors. These factors are mobilized to solve problems in reading according to the readers' purposes and the demands of the task stimuli.
1957	Chomsky	*Syntactic Structures.* Differentiates competence and performance. Kernel sentences generated by phrase structure rules; transformation rules operate on syntactic structures to express different types of sentences. Attacks S-R explanations of language acquisition.

1960	Ausubel	Advance organizers are abstract knowledge structures which assimilate information in text.
1966	Goodman	Reading consists of psycholinguistically determined expectancies or hypotheses used by the reader and the text, and sampling of print to test and confirm hypotheses.
1970	Singer & Ruddell	*Theoretical Models and Processes of Reading*. Perceptual, linguistic and cognitive processes explain reading.
1972	Bransford & Franks	Readers chunk semantically related information across sentences.
1972	Winograd	"World Knowledge," information stored in long term memory, is necessary for comprehending texts.
1974	Kintsch	*The Representation of Meaning in Memory*. Meaning is represented by given and inferred propositions.
1975	Meyer	*The Organization of Prose and Its Effects on Memory*. Texts are represented by hierarchically organized propositions stored in memory.
1975	Rumelhart	A grammar for a well organized story, such as a fable, consists of a set of recursive rules that include plot, character, problem, situation, and resolution.
1976	Gough	Information processing model accounts for psychological data on reading.
1976	Rumelhart	Reading consists of an interaction between text-based data and reader-based resources.
1977	Schank & Abelson	Scripts, sequences of events stored in long term memory, are necessary for explicating and assimilating texts and culturally determined situations ("Mops" have replaced scripts).
1977	LaBerge & Samuels	Model of automaticity in processing print describes skilled readers' performance.
1977	Anderson, Spiro, Montague	*Schooling and the Acquisition of Knowledge*. Role of schemata, defined as knowledge structures with variable slots, explains assimilation of text data, construction of text interpretation, and inferential reconstruction upon recall.
1977	Van Dijk & Kintsch	Texts are represented in memory by macrostructures and microstructures.
1979 & 1980	Tierney & Spiro; Tierney & LaZansky	Comprehension of a text consists of a band of acceptable interpretations that vary on a continuum of text-based and reader-based interactions.

At the turn of the century, reading was a central topic in psychology. Javal (1879) had already discovered and measured eye-movement behavior in reading. Wundt's first American doctoral candidate, Cattell (1886), had found that adults perceived a word more rapidly than its individual letters. Huey (1908) proclaimed in his *Psychology and Pedagogy of Reading* that the understanding of reading would be the acme of a psychologist's achievement. In 1917, Thorndike defined reading as reasoning. Although his definition changed the goal of reading, it may have also deterred some psychologists from studying reading comprehension as such.

With the advent of Watson's behaviorism, psychological studies of reading focussed on eye-movement behavior in reading. Buswell (1922) recorded developmental changes over grades 1 to 12 in eye-movement behavior during reading. Judd and Buswell (1922) demonstrated that, during silent reading, readers varied their eye-movement behavior according to the kind and difficulty of material, and their purposes in reading. Although some research was done on reading in the content areas during this time (Bond, 1938; Strang, 1942;

Robinson and Hall, 1941), psychologists limited their work mostly to visual perception and eye-movement behavior in reading. Consequently, Anderson and Dearborn's book on the psychology of teaching reading, published in 1952, consisted mostly of research on visual perception, eye-movement behavior in reading, Olson's organismic age concept for determining expectancy in reading achievement, and S-R explanations of methods of teaching reading. This research was atheoretical and contained hardly any studies on comprehension. Indeed, the prevailing conception was that reading consisted of word perception; all else was thinking (Buswell, 1952). This conception was supported by Thurstone's (1946) factor analysis of reading tests into two factors: word meaning and reasoning.

Two notable exceptions to S-R type research occurred during the behavioristic era. Influenced by gestalt psychology, Bartlett (1932), an English psychologist, studied the dynamics of memory. He had readers recall an Indian story, entitled "The War of the Ghosts." Bartlett found that readers constructed an interpretation of the story, and, over time, their recall of the story occurred as an inferential reconstructive process based upon their schemata. Some 40 years later, when the zeitgeist had changed to favor cognitive psychology, American psychologists were to hail Bartlett's research as a pioneering investigation.

The other exception was the multiple regression studies conducted by Holmes (1948, 1953, 1954) to test the central hypothesis of his substrata factor theory of reading. He found that a multiplicity of factors accounted for individual differences in general reading ability and for predictors of this ability. Holmes theorized that these factors functioned at various substrata levels, but were mobilized and organized into momentary working systems according to the reader's changing purposes and the demands of the reading task.

During the 1950s, the dominant experimental paradigm for studying verbal learning was paired-associates. "Reading," which had been the title of a chapter in Woodworth's (1938) book *Experimental Psychology*, was reduced to only a topic in his 1954 revision (Woodworth and Schlosberg, 1954). This narrow scope of research in reading changed dramatically after Chomsky (1957) wrought his revolution in linguistics and after his severe attack on S-R psychology (Chomsky, 1959). Apparently influenced by Chomsky's arguments, many pschologists abandoned the paired-associates paradigm with its S-R explanations and shifted their research to sentence comprehension with cognitive explanations. Then, research in psycholinguistics began to flourish and greatly affected the field of reading. Goodman (1966), for example, drew upon expectancy theory and psycholinguistics to explain the process of reading.

However, psychologists soon discovered that Chomsky's linguistic theory presupposed rather than predicted meaning. So they changed their research to focus on meaning (Van Dijk and Kintsch, 1977). Ausubel (1960) had already formulated his theory of comprehension, which features advance organizers—abstract knowledge structures that assimilate information lower in the knowledge hierarchy, and can be modified to accommodate new information. Bransford and Franks (1974) discovered that syntax was necessary to get meaning correctly into long-term memory where it was chunked across semantically related sentences.

Some psychologists argued that meanings were stored in memory in the form of propositions, and began to use propositional analysis to represent the text-based input (Kintsch, 1974; Meyer, 1975; Frederiksen, 1975). They also began to investigate the effects of text variation on reader's speed of processing text and on the reader's construction of macrostructures and microstructures (Kintsch, 1980; Meyer et al., 1980). They have assumed that propositional meaning is sufficient for comprehension of a text. But they have not yet determined normative levels of inference nor the developmental conditions under which these inferences are made (Singer, 1980a).

Another source of impact on reading came from the use of computers. During the 1960s, computers became readily available and began to be used for studying artificial intelligence, including simulation of reading comprehension. Winograd (1972) discovered that he had to program into computers "world knowledge" to simulate the information readers had accumulated in their long-term memories and had to mobilize in order to explicate and comprehend printed materials. Schank and Abelson (1977) added the concept of "scripts" which represent the common knowledge for sequences of events that readers acquire and mobilize for guiding their behavior and interpreting situations. (I understand Schank has substituted for "scripts" the concept of "mops" which are memory organization points.)

A boon to research in reading comprehension came from another source. In the 1970s, federal funds began to emphasize research on reading comprehension. The National Institute of Education established a Center for the Study of Reading at the University of Illinois under the very able direction of Richard Anderson. The Center has been productive in studies on the role of schemata for assimilating, interpreting, and comprehending printed materials. This view of reading comprehension is presented in the book *Schooling and the Acquisition of Knowledge* (Anderson, Spiro, and Montague, 1977).

Many of the new concepts in reading were incorporated into Rumelhart's (1976) interaction model of reading. After criticizing LaBerge and Samuel's (1976) model on automaticity and Gough's (1976) model on information processing, Rumelhart presented his interaction model. In this model, readers allocate attention between text-based data and reader-based resources. The resulting information is integrated and funneled through the reader's limited capacity processor. A corollary to the interaction model is that reading comprehension consists of a band of acceptable interpretations that vary on a continuum of text-based and reader-based interactions (Tierney and Spiro, 1979; Tierney and LaZansky, 1980). These interpretations are represented in memory by macrostructures and microstructures (Van Dijk and Kintsch, 1977).

Thus, in the last 20 years, psychology has restored reading to a central position. This renewed status of reading is exemplified by such books as *Basic Studies on Reading* (Levin and Williams, 1970), *Language by Ear and by Eye* (Kavanagh and Mattingly, 1972), *The Psychology of Reading* (Levin and Gibson, 1975), *Theoretical Models and Processes of Reading* (Singer and Ruddell, 1970, 1976), *Basic Processes in Reading: Perception and Comprehension* (LaBerge and Samuels, 1977), *New Directions in Discourse Processing* (Freedle, 1979), and *Perception of Print* (Tzeng and Singer, 1981).

But the explosion of knowledge in the field of reading has primarily been on the psychology, not on the pedagogy of reading. Although some researchers, such as Dave Pearson and Rob Tierney at the Center for the Study of Reading, are conducting research on teaching comprehension, the gap between knowledge of reading and application of this knowledge to instruction for the development and improvement of reading comprehension is great and is growing greater. We need to close this gap by testing implications derived from psychological theories and research in classroom settings. We need to determine whether these implications are developmentally valid, and if so, whether they can be translated into instructional strategies that can improve reading comprehension in normal classroom settings. I shall summarize a list of some research-based hypotheses and implications that need to be tested and validated at the classroom level. The list is shown in Table 2.

TABLE 2

List of Research-Based Hypotheses for Testing and
Devising Strategies for Improving Comprehension
at the Classroom Level

1. *Input stimuli.* Modify input materials so that they fit readers' abilities; then develop readers to higher levels of ability on the relevant variable.
    a. *Shift conceptual levels from abstract towards concrete*
        (1) *word meaning:* substitute high for low frequency words (Marks, Doctorow, and Wittrock, 1974).
        (2) *sentence comprehension:* concretize sentence (Anderson, 1976); this modification reduces sentences from more abstract levels and may enable readers to then generate instantiations and consequently comprehend the sentences. In some cases, sentences may be reduced from general to particular statements that tend to make them more imageable and hence more memorable (Levin, 1976). For example, compare these sentences: The regulations annoyed the salesman vs. The *parking* regulations annoyed the salesman.
        (3) *passage sequence:* construct concrete passages to develop advance organizers that precede more abstract passages (Royer and Cable, 1975).
    b. *Modify syntax*
        (1) *Syntactic level:* use syntax in print that corresponds with the syntax employed by readers in their oral language (Ruddell, 1965).
        (2) *Syntactic complexity:* combine sentences to explicate their causal relationships, which will make the sentences longer. But, in contrast to the readability hypothesis which asserts that shorter sentences are easier to comprehend, some longer sentences with the causal relationship explicated will be easier to comprehend (Pearson, 1976). For example, compare these sentences: Mary went to the store. She bought a loaf of bread vs. Mary went to the store *to* buy a loaf of bread.
        (3) *Embedded clauses:* whenever choice can be exercised, use right branching (clauses embedded after the main verb) to reduce memory load on short-term storage and hasten storage in long-term memory (Gough, 1976).
    c. *Verb voice*
    For ease in processing print, use the active over the passive voice, but when the passive has to be employed, do so with agent expressed (Wanat, 1976).
    d. *Text organization*
        (1) *paragraph organization:* explicate situational contexts close to the beginning of a passage so that a global representation can be attained early in the passage and used for interpreting events (Bransford and Johnson, 1972).
        (2) *clarity:* explicate macrostructure and microstructure relationships (Van Dijk and Kintsch, 1977).
        (3) *density:* increase redundancy in text so that a higher proportion of text can be relevant to a particular instructional objective (Rothkopf and Kaplan, 1972).
        (4) *information:* explicate information needed for making inferences (Marshall, 1975).

        (5) *signaling terms:* include these terms in texts (Meyer et al., 1980; Robinson, 1975) and teach students to use them.

2. *Mediational processes.* Use instructional procedures that promote development of cognitive processes for encoding, storing, and retrieving information.
    a. *Questioning strategies*
        (1) *direct questioning:* question students on those passages they are to learn, remember, and recall on a comprehension test (Anderson and Biddle, 1975).
        (2) *objectives or goals:* explicitly state specific objectives, purposes, or goals that are related to statements in the text (Rothkopf and Kaplan, 1972).
        (3) *self-questioning and self-selection of goals:* have students generate their own questions (Manzo, 1970; Frase and Schwartz, 1975; Singer, 1978) and formulate their own reading goals. Probably self-formulation of goals will have benefits similar to the advantages of self-questioning; used in combination, the interaction of the two processes may even result in a greater degree of comprehension, or at least in more efficient attainment of comprehension.
    b. *Verbalization.* Teach students to summarize and paraphrase text (Wittrock et al., 1975).
    c. *Knowledge*
        (1) *Structure of knowledge:* teach students a wide range of organized information through a broadly based liberal arts curriculum.
        (2) *Mental encyclopedia:* develop a broad repertoire of information through direct and vicarious experience (Winograd, 1972). Teach students to use their background information (Hansen and Pearson, 1980).
    d. *Word meanings, concepts, and schemata:* have students develop within the limits of their increasing capacities a lexicon that has range, altitude, and depth (Russell, 1956), and organize it into a hierarchical structure (Rumelhart, 1975; Anderson et al., 1977).
    e. *Morphemes:* teach students to analyze printed words into their morphemic constituents (Ruddell, 1976; Singer, 1976).
    f. *Imaging:* teach and encourage students to transform meanings expressed in print into images (Levin, 1976, 1977).
    g. *Purposes*
        (1) counsel students to be initiators, that is, establish their own comprehension goals (deCharms, 1972).
        (2) develop students' self-confidence and independence (Athey, 1976).
    h. *Reasoning about print*
        (1) teach students to analyze sentences into given and new information and compare it with information already in their mental encyclopedias (Clark and Clark, 1977).
        (2) have students learn and practice generation of concrete scenarios from their abstract structures; have students also generate examples or hypotheses for superordinate schemata (Anderson et al., 1976).
        (3) teach students to form new schemata by having them abstract and generalize from direct and vicarious experiences using a teaching process such as Taba's questioning sequence (Taba, 1965).
3. *Output Products:* use the type of measurement that corresponds with the reader's purpose, mode of processing, and demands of the material.
    a. *Letter-by-letter:* reading that follows more closely to the sensory-driven level of the processing continuum is more likely to occur when the exact meaning of the author is sought. At the extreme level, letter-by-letter processing (Gough, 1976) or direct perception (McConkie and Rayner, 1976) occurs, but does not actually involve processing all the letters, only those needed to identify the word (Strange, 1979). For this mode of processing and reader goal, verbatim scoring of a recall protocol is appropriate.
    b. *Hypothesis and sampling of print:* a mode of reading that is closer to the

concept-driven level is more likely to occur when (a) the reader only wants to extract the general meaning of the passage or (b) the material is familiar or (c) the material only demands extraction of a general idea. Then the reader is more likely to hypothesize, sample the print, and confirm or disconfirm hypotheses (Goodman, 1976; Wanat, 1976). The same type of processing may occur when the contextual situation is missing or has not been stated yet and the reader is testing out ways of conceptualizing the input and consequently shifts from one hypothesis to another (Rumelhart, 1978). For this type of reading, gist or lax scoring would be more relevant than verbatim or exact scoring (Nicholson, 1977).

c. *Level of Macrostructures and Microstructures:* use text bases that specify macrostructures and microstructures which are appropriate to students' developmental levels. Teach students to form these structures, and supplement them with strategies for processing various types of expository text (Kintsch, 1980; Singer, 1980a).

d. *Coherence in Teaching and Testing:* make sure that the readers' purposes, information in text and ways of processing them, and goals or tests cohere. Or stated in another way, determine whether the directions and tasks given the reader fit the information provided in texts and the questions asked on comprehension tests (Rothkopf, 1980; Singer, 1980b).

HYPOTHESES FOR CLASSROOM TESTING

The first set of hypotheses deals with the modification of input stimuli to fit readers' abilities. These hypotheses refer to levels of abstraction, syntactic complexity, and organization in text, including discourse and rhetorical devices, lexical choices, and explication of propositions. Such modification may lead to the determination that an input variable is related to comprehension. After this determination has been made, it is necessary to try to improve the process or knowledge entailed by this variable so that readers can comprehend the more difficult and more varied material. This type of improvement is necessary if the claim is made that variation in the input stimuli has improved comprehension.

The second set of hypotheses focuses on improvement of mediational processes. These hypotheses refer to acquisition of self-questioning and verbalization strategies, development of information and knowledge structures and their use in reading, improvement of students' lexicons and morphological ability, instruction in use of imagery in processing print, and emphasis on teaching students to vary their purposes in reading and increase their ability to reason about print.

The last set of hypotheses refers to the assessment of comprehension. The main point in these hypotheses is that the type of assessment used should correspond with the reader's purposes, mode of processing, and level of reasoning abilities. In general, there should be coherence among the reader's task, the information in the text, the test for assessing comprehension, and the teacher or adjunct aids that affect each of these components for reading and learning from text (Rothkopf, 1980; Singer, 1980b).

In testing these hypotheses and in teaching for the improvement of comprehension, we have to observe some principles and avoid some pitfalls in reaching conclusions based on classroom instruction. These principles and pitfalls are summarized in Table 3.

TABLE 3
Principles and Pitfalls in Testing Hypotheses on Teaching Reading Comprehension at the Classroom Level

1. *Phase out the Teacher and Phase In the Student While Teaching a Process of Comprehension.*
   Example: Active Comprehension: From Answering to Asking Questions (Singer, 1978, 1979)
2. *Transfer of Training Test:*
   Negative Example: Wittrock, Marks, and Doctorow (1975)
   Comprehension strategy of summarizing paragraphs—not tested in a transfer paradigm.
3. *Developmental Validity of Hypotheses:*
   Example: Instantiation hypotheses, confirmed at the college level (Anderson et al., 1976) did not generalize to grade school level (Dreher, 1980).
4. *Develop schemata in a broadly based curriculum which stresses development of vocabulary, concepts, and range of information:*
   Example: Prediction that, all else being equal, narrowing the curriculum is likely to reduce reading comprehension.
5. *Devise and Test Strategies for Development of Schemata: Example:* Hayas (1979) used baseball analogies to develop schemata for comprehending cricket.
6. *Teach and Test for Improvement of Cognitive Processes: Example:* Hansen and Pearson (1980) devised prediction strategy for teaching second graders to use spontaneously their background knowledge for making inferences from text.
7. *Ascertain that materials of appropriate reading difficulty are used for testing hypotheses on both the good and poor readers:*
   Example: Taylor (1979) inferred that fifth graders whose reading ability was at a third-fourth grade level could use knowledge-based or reader-based strategies when reading familiar material, but were reduced to text-based strategies for unfamiliar material.
8. *Use downward adaptation of materials to fit readers' abilities to test relevance of a particular variable; then instructionally intervene to develop that variable in students and determine whether comprehension improves as a result:*
   Negative example: Ruddell's (1965) match of sentence syntax in texts to students' oral syntax.
9. *Use an assessment measure that is appropriate and sensitive to the research or instructional strategy:*
   Examples:
   a. Tierney and Spiro (1979) and Tierney and LaZansky (1980): Band of acceptable responses for continuum of text-based and reader-based interactions.
   b. Level of inference for macrostructures and microstructures have to be developmentally appropriate (Singer, 1980a).
   c. Direct perception (McConkie and Rayner, 1976); Gough, 1976) occurs when the reader's purpose is to be precise. Scoring can be verbatim.
   d. Hypothesis or expectancy sampling and confirmation (Goodman, 1966; 1976) occurs when the material is familiar and/or reader's purpose only requires construction of a general idea (Nicholson, 1977).
   e. Ethnic texts assessed by scriptural comprehension show comprehension advantage for ethnic minorities (Rogers-Zegarra, 1980).
   f. Probe after free recall to test for (a) treatment effects (Hansen and Pearson, 1980); (b) perspectives (Pichert and Anderson, 1977; Anderson and Pichert, 1977).
   g. Use tests that have curricular and instructional validity. They should also have research validity, that is, be sensitive to experimental

treatment (Katz, 1980). In short, provide for coherence among readers' purposes, text information, and criterion referenced goals, and the effect of teacher or adjunct aids on these components of reading and learning from text (Rothkopf, 1980; Singer, 1980b).

10. *The instructional fallacy is the assumption that if students do not learn, the difficulty lies with the students, not the method or the instruction. To try to avoid this fallacy, during an experiment, test to determine the relevance and effectiveness of experimental or instructional conditions for the learner's aptitudes, abilities, and motivation.*[2]

PRINCIPLES AND PITFALLS IN TESTING HYPOTHESES ON READING COMPREHENSION AT THE CLASSROOM LEVEL

Phase Out the Teacher and Phase in the Student While Teaching a Process of Comprehension

A multiplicity of strategies has been constructed to facilitate attainment of comprehension. The Directed Reading Activity (DRA), which is used in most basal readers, was the first strategy to gain widespread use. Others are Robinsons' SQ3R (1946, 1961), Ausubel's advance organizer (1960), Herber's reading guides (1970), Earle and Barron's overview guides (1973), and marginal glosses (Singer and Donlan, 1980). Rothkopf's (1966, 1980) research on adjunct aids, consisting of directions to the reader, questions embedded in text, and selection of reading goals in the form of relevant test questions has provided considerable knowledge on teaching students to comprehend printed materials, as has Frase's (1967, 1970) work on the effect of questions, text organization, and document design. Richard Anderson and W. E. Biddle (1975) have theorized that teacher-posed questions serve to transfer information from short-term to long-term memory. Recently Tom Anderson and Bonnie Armbruster (1980) reported that a study technique such as outlining or underlining, is only effective to the extent that it leads the reader to concentrate time and effort on the reading task.

However, the goal of instruction is reader independence. In other words, instruction should teach students to utilize strategies on their own. For example, teachers should aim to teach students to formulate and read to answer their own questions.

I have referred to self-questioning as a process of active comprehension, which is a continuous process of asking and answering questions during reading (Singer, 1978, 1979). To attain this goal, we (Singer and Donlan, 1980) have advocated a phase-out/phase-in strategy. In this strategy, the teacher takes the first step by modeling questions that are appropriate to the content. The teacher also begins to develop students' knowledge of the content. We know that each content area has questions that are specifically appropriate to it (Broudy, 1977) and that knowledge of content is necessary for appropriate question-generation (Miyake and Norman, 1979). Furthermore, text-types, such as expository prose, have their own question strategies, frequently cued by signaling words (Smith, 1964a, 1964b; Robinson, 1975; Meyer et al., 1980). As students acquire some knowledge of the content, the teacher takes the next step of forming them into groups to ask each other questions. Eventually the teacher can take the final step of having them ask and answer appropriate questions on their own.

Can students learn to generate their own questions? Rosenthal, Zimmerman, and Durning (1970) used comprehension of pictures to demonstrate that sixth graders could learn to imitate their teacher's questions. But the students did not simply copy their teacher's questions, as shown by the student's ability to transfer their learning to asking questions for comprehending new pictures. However, the investigators did not determine whether improvement in self-questioning transferred to comprehension of text.

Frase and Schwartz (1975), however, did investigate the use of a self-questioning technique for comprehending a text. They reported that 48 students who studied by asking each other questions, mostly verbatim questions, attained significantly higher recall scores than students who studied alone. The experimenters also found that college students attained higher scores on text pages for which they generated questions than on alternate pages which they only studied.

Self-questioning for learning from text also works at the elementary level. Conducting their research at the fifth grade level, Helfeldt and Lalik (1976) found that the reciprocal questioning group performed significantly better on a Van Wagenen interpretation test than a unilateral, teacher-posed question group. The reciprocal question group was allowed to ask the teacher a question for each teacher-posed question they answered.

Self-questioning was also superior in a remedial reading program for students whose ages ranged from 7 to 25. In this program, Manzo (1970) had trained teachers tutor students through use of a technique called "ReQuest." The technique consists of the teacher requesting students to ask questions. The results indicated that the "ReQuest" group performed significantly better on two standardized comprehension tests than a control group which had been taught by the Directed Reading Activity which features teacher-posed questions.

It is possible that good readers at the high school level have learned self-questioning techniques on their own. This interpretation would explain why Andre and Anderson (1978-1979) found that training high school students in self-questioning was more effective with the lower than with the higher verbal ability students.

An hypothesis to explain the effects that student-posed questions have upon learning and comprehension is that self-questions focus attention and emphasize text passages that answer these questions. The answers, stored in long term memory, are then retrieved by test questions that contain verbatim cues. But, unlike teacher-posed questions, student-posed questions can be asked continuously throughout the text. Consequently, students are more likely to perform better on the criterion test because they have asked and answered more questions as they read. This explanation assumes that teacher modeling was used to teach the students which questions were relevant to the content and that the students had acquired some knowledge of the content prior to generating their own questions. Under these conditions, active comprehension is likely to lead to more review, more mathemagenic behavior, and more storage in long-term memory of information gained from text.

To Claim that Learning Has Occurred, Test for Transfer of This Learning to New Passages

Another principle is that a transfer test should be used to determine whether students have learned a strategy or a process for improving their comprehension. Otherwise, the conclusion of a classroom experiment should not refer to what the experimental group did or did not learn. We shall review a research study to illustrate this principle.

Wittrock has theorized that reading is a two-stage process. In the first stage, the reader uses input data, including the semantic, syntactic, phonetic, and episodic characteristics of words in a specific context to reduce uncertainty in the identification of previous experiences stored in long-term memory. In the second stage, the reader generates or constructs from the identified previous experience one or more distinctive representations of an event or situation consistent with the words of the sentence. This actively constructed representation, induced from memory of prior events, is the psychological meaning of the sentence. Testing an hypothesis from this theory, Doctorow, Wittrock, and Marks (1978) asked fifth and sixth grade students to generate sentences that would summarize the meaning of paragraphs in a story. They found that these students doubled their comprehension on the criterion-referenced test that accompanied the story they had summarized. But the experimenters did not attempt to determine whether the students could transfer the summarization strategy to another story without experimenter guidance.

Another principle we should apply before we use a psychological hypothesis as a basis for devising an instructional strategy to improve reading comprehension is this:

Determine the Developmental Validity of the Hypothesis

Hypotheses are frequently validated on college samples. But the developmental generality of the hypothesis also has to be tested on an elementary sample before basing instruction upon it. For example, Anderson et al. (1976) validated the instantiation hypothesis on a college sample by demonstrating that these students substituted and stored in memory an instantiated term for a general term. They found for the sentence, "The woman was outstanding in the theater," that the instantiated term "actress," was a better retrieval cue for the object noun of the sentence, "theater," than was "woman," the general term that had actually occurred in the sentence. Dreher (1980) tested the generality of this hypothesis at the fifth and eighth grades. In the first part of the experiment she requested the students to generate for each sentence a particular term for the general term; the students were able to comply. But, under the direction to read and remember the sentences, another sample of students did not *spontaneously* instantiate the sentences.

Whether students can learn to instantiate spontaneously remains to be seen. In the meantime, children's lack of spontaneity in this process is consistent with Brown's (1977) theory on developmental changes in cognitive or executive control. Apparently the development of executive control for instantiation does not occur until sometime after the eighth grade.

What does increase developmentally and cumulatively is the knowledge students gain from their curriculum experiences. This fact leads to this principle for improving reading comprehension:

To Improve Comprehension, Use a Broadly Based Curriculum which Intentionally and Systematically Develops Vocabulary, Information, and Schemata

Are schemata currently being developed and organized hierarchically in students? They are, but not exclusively in instruction on comprehension. They are mostly being developed in the content areas of the curriculum, but probably not intentionally nor in a hierarchical organization. Nevertheless, we would expect that a person who had been educated in a broadly based curriculum, all other things being equal, would have better comprehension over a greater range of content than a person who had been educated in a narrower curriculum. The more broadly educated individual would have gained a wider range and depth of information. Thus, we conceive of the entire curriculum, including all the direct and vicarious experiences that students gain in and out of school, as contributing to the range, depth, and altitude of an individual's schema hierarchy and knowledge of the world. These schemata, stored in the reader's mental encyclopedia, enable readers to explicate passages, make inferences, and assimilate information. They are a major contributor to the individual's reading comprehension.

Conversely, we would predict that when schools narrow their curriculum, students' comprehension is likely to go down. The San Diego Unified School District, in an attempt to satisfy a court order to raise reading achievement from the 30th to the 50th percentile rank in some 17 inner city schools over the next three years, is complying by increasing instruction in reading, math, and language, and decreasing it in other areas of the curriculum, such as science, history, and geography. We predict that this narrowing of the curriculum is likely to lead to a decrement, not an improvement in reading comprehension. Consequently, we suggested to San Diego's director of research and evaluation that San Diego conduct an experimental test of this prediction by randomly assigning half of the schools to the narrower curriculum and the other half to the present or perhaps to a broader curriculum. At this time, we do not know whether this naturalistic type of experiment will be conducted.

If our prediction is confirmed, we would attribute it to (1) acquisition of new schemata as a result of instruction and learning in various content areas and (2) use of new and previously acquired schemata for assimilation of information and construction of meaning during reading and for inferential reconstruction upon recall.

Although the accommodation process for the acquisition of new schemata is germane to education in general, we know very little about the process. For example, we need to know whether we can apply the accommodation process in classroom situations. If so, what strategies result in acquisition of schemata and under what conditions? Therefore, we recommend this principle for those researchers concerned with the improvement of reading comprehension:

To Conduct Research on the Improvement of Reading Comprehension, Devise Strategies for Development of Schemata

Hayes (1979), who completed a doctoral dissertation under the direction of Rob Tierney, conducted a study that fits this principle. Using baseball analogies for learning schemata about cricket, he found that the analogies affected recall of the text in different ways, depending on when the analogies were given. Analogies given in advance led to more recall of explicitly stated text-based information. Analogies embedded in text resulted in more reader-based information and more schema generalization as shown by transfer of learning to prediction of certain cricket match situations. Analogies given in advance *and* embedded in text led to greater schema specialization as shown by ability to discriminate between cricket and non-cricket information.

Hayes's findings have great relevance for construction of texts and for knowledge of the accommodation process. The evidence is clear that analogies facilitate learning from text. Whether readers can learn to use an analogy even more effectively and whether they can improve their learning from texts without advance or embedded analogies by learning to generate their own analogies are questions that still need to be answered. In short, researchers in education need to determine whether teachers can improve students' cognitive processes for attaining comprehension. To do so, researchers have to create instructional strategies, teach them to students, and assess their effect upon comprehension. The principle that fits this requirement for improving comprehension is:

Teach and Test for Improvement of Cognitive Processes for Encoding, Storing, and Retrieving Information

Operating according to this principle, Hansen and Pearson (1980) devised a prediction strategy to teach second graders to draw spontaneously upon their background knowledge in order to make inferences from text. The strategy asked the children to answer a question about their previous experiences related to ideas in the story and then to hypothesize or make predictions about something similar that might happen in the story. The experimenters hoped the strategy would teach the students to consciously try to integrate text and prior knowledge during reading. Another experimental group was given practice in answering inferential questions. Hansen and Pearson expected the practice either to improve the children's processes of inference or else to develop anticipation of questions which would lead them to focus on interpretation rather than on a literal level of comprehension. The results indicated no significant differences on a recall measure, but both experimental groups were superior to the control group on the 10 item open-ended literal comprehension questions and on the 10 item, open-ended inferential comprehension questions which required use of prior knowledge. Hansen and Pearson also found that on the *Stanford Achievement Test,* which requires inferential reasoning, the question group was superior. Although the experimenters demonstrated that their teaching strategies improved second graders' ability to integrate text with prior knowledge and make valid inferences, they did not test whether the strategy transferred to stories on which the students had no pretraining. However, the Question Method apparently did lead to generalized improvement in making inferences as assessed by the *Stanford Reading Achievement Test.*

Hansen and Pearson used the experimental-control group design. This paradigm does not lead to a pitfall sometimes encountered when the good vs. poor reader contrast is employed. In the latter experimental design, caution should be exercised lest an implication of deficit is drawn for the poor readers. The principle to follow is this:

Ascertain that Materials of Appropriate Reading Difficulty are used for Testing an Hypothesis on both the High and Low Reading Groups

This principle was violated in a recent study by Meyer, Brandt, and Bluth (1980). To test for the effect that structural signal words in a passage have upon recall of propositions, they gave a reading passage with and without structural signal words to good and poor ninth grade readers. The results revealed that when the signal words were available in the text, the good readers used them to organize their recall but the poor readers did not do so; instead, they simply gave a list of propositions. But Meyer, Brandt, and Bluth did not give the poor readers a passage appropriate to their reading level to see whether they too could then use structural signal words to organize their recall. Instead, they implied that the poor readers had a deficit in their ability to organize recall of the text.

Downward adaptation of the text to the ability level of the students, however, should not be interpreted as an improvement in comprehension. Such a downward adaptation may be a way of determining whether a particular variable is adversely affecting a group's comprehension. The next step is to provide instructional intervention and then test for improvement in comprehension. The principle is:

Use Downward Adaptation of Material to Fit Readers' Abilities and to Test the Relevance of a Particular Variable for Comprehension; then Instructionally Intervene to Develop that Variable in Students and Determine Whether Comprehension Improves as a Result

An example of downward adaptation is Ruddell's (1965) matching of sentence syntax in printed materials to fourth graders' oral sentence syntax. The match resulted in better performance on a cloze text of comprehension. However, Ruddell did not go on to the next step to improve the fourth graders' use of syntax in comprehending printed materials. Consequently, he could not and did not attempt to claim that he had improved the students' general comprehension.

The cloze test was a sensitive and appropriate test to use in Ruddell's experiment. Indeed, a relevant test is necessary for any experiment or for assessment of any instructional effect. The principle is:

Use an Assessment Measure that is Appropriate and Sensitive to the Research or Instructional Strategy

Although we may be able to agree on this principle, carrying it out in practice is difficult, even more so because of the interaction concept of reading comprehension, and because we are aware of the role of schemata and inference in construction of meaning during reading and inferential reconstruction upon recall. For example, Tierney and Spiro (1979) and Tierney and LaZansky (1980) have pointed out that comprehension consists of a band of acceptable responses that vary on an interaction continuum from text-based to reader-based. Which end of the continuum is dominant depends on the readers' purposes, abilities, experiences, and the demands of the material. Moreover, for evaluating comprehension, we have to determine the level of macrostructures or inferences for a text-base that is developmentally appropriate (Singer, 1980).

We also have to decide, for any comprehension measure, whether to use verbatim or lax scoring. We do have some criteria for solving this problem. If the reader's purpose is to get a precise meaning for a text, then McConkie and Rayner's (1976) direct perception hypothesis is tenable; so is Gough's (1976) position that readers "plow" through each word until the word is identified (Strange, 1979). For this reading purpose, verbatim scoring would be appropriate.

But, if the reader's purpose is to attain only the general meaning of a passage, if the material is familiar, and if the demands of the material require construction and retention of only a general idea, then Goodman's (1976) position is defensible. His position is that readers hypothesize, anticipate meaning, sample print, and confirm or disconfirm hypotheses as they read. For this position, lax or gist scoring would be suitable (Nicholson, 1977).

Instructional conditions that move attainment of comprehension towards the reader-based end of the interaction continuum occur when ethnic texts are used and scriptal comprehension is assessed. Under these conditions, Rogers-Zegarra (1980) has found that the Mexican-American group scored higher than the Anglo group for the Mexican-American material and equalled the Anglos on the Anglo materials.

Whatever input text is used, we cannot rely upon free recall alone to find out whether groups differ; we must also probe, as Hansen and Pearson (1980) found they had to do in their study. We have to ask not only for literal recall, but also for text-based inferences and answers to reader-based scriptal questions, including questions on different perspectives for a text (Pichert and Anderson, 1977).

Another factor that affects the comprehension output is the time delay between the reading of a text and administration of a comprehension test. If we test immediately after the reading of a passage, we are more likely to get a reproduction of the text; but if we allow time to intervene before we test, we are likely to get more reconstructive intrusions.

In general we should use tests that provide for instructional validity, that is, there should be coherence among readers' purposes, text information, and criterion-referenced goals, and the effect of teacher or adjunct aids on these components of reading and learning from text (Rothkopf, 1980; Singer, 1980b).

Tests should also have research validity, that is, the tests used in a study should be sensitive to experimental treatments. For example, in retesting the hypothesis of the First Grade Studies which have been reported by Bond and Dykstra (1968), Katz (1980) constructed criterion-referenced tests for each method of supplementary instruction (graphophonemics, morphophonemics, semantics, and syntax) given to her four experimental groups. She found that the supplementary instruction did make a difference on the criterion tests and in the multiple regression equation for predicting comprehension. The variation in supplementary instruction had no differential effect on comprehension, but, in contrast to the First Grade Studies's conclusion, it did affect development of children's subsystems for attaining comprehension.

Because of the choices involved in assessing comprehension and because the particular assessment instruments used affect the results of an experiment so much, we should follow this principle in planning and publishing research on comprehension:

Use a Comprehension Measure that is Appropriate to the Condition of the Experiment or Instruction, and State the Rationale for the Test in Publishing Research Results

Finally, we come to our last pitfall in teaching comprehension, the instructional fallacy. This fallacy assumes that if students do not learn, the difficulty lies with the students, not with the method or the instruction.[2] But as Carroll (1963) has specified in his conditions of school learning, we have to consider the outcome of schooling as a function of variables in the student and variables in the instructional or experimental situation. Hence, in addition to consideration of statistical power tests (Levin, 1975), validity of experimental designs (Campbell and Stanley, 1963), and robustness of treatments or instructional strategies, we also have to consider the instructional fallacy in evaluating classroom research.

We can try to avoid this pitfall by applying this principle:

During the Experiment, Test to Determine the Relevance and Effectiveness of Experimental or Instructional Conditions for the Learner's Aptitudes, Abilities, and Motivations. Then, in Reporting, State the Instructional Conditions for Both the Experimental and Control Groups, including the Rationale for the Criterion Test

Space permits me to provide only one example of what I mean by a test of experimental conditions. This test is one way to answer a perennial question: How long should instruction last until a criterion test is given? Borrowing from conditions for paired associate learning, Dan Donlan and I (1979) found that one answer to the question is to alternate teaching and test conditions. For example, we wanted to determine whether students could learn to use a problem solving schema for generating their own questions in comprehending short stories. For each story, we provided teacher-modeling of schema-based questions followed by criterion-referenced tests and plotted the difference each time between our experimental and control group. We reasoned that if the treatment was effective and students were learning to generate their own schema-based questions, the differences would accumulate. That is what we found. After five stories, we obtained a clear separation between our experimental and control groups on the criterion test.

SUMMARY AND CONCLUSIONS

Scientific revolutions in psychology have resulted in dramatic changes in research in reading. After almost a century of relatively few landmark studies in the psychological investigation of reading, an explosion of knowledge on reading acquisition and learning from text has occurred in the last 20 years.

However, the resulting hypotheses at the input, mediational, and output stages of reading still need to be tested at various school levels to determine their developmental validity. Instructional strategies then need to be created and tested to determine whether they can improve comprehension in ecologically valid settings.

This research is fraught with numerous pitfalls but the task of exploring and improving the mental cosmos from which reading comprehension emerges should be a challenge to competent and creative researchers. They will have to devise the strategies, avoid the pitfalls, follow certain principles for improving comprehension, and conduct research in classroom settings to determine the validity of the instructional strategies. We urge educational psychologists to concentrate their efforts on this task, the National Reading Conference to highlight these studies, educational journals to feature reports of this research, and funding agencies to support it. Together, these agencies can make the 1980s an era for research on instruction and improvement of reading comprehension.

ACKNOWLEDGEMENT

Appreciation is expressed to Marian Jean Dreher, now at the University of Maryland, Olga Kokino and Betty Medved for their assistance on this paper.

I am grateful to John Guthrie and Frank Murray for their discussion with me about the instructional fallacy (p. 15-16).

REFERENCES

Anderson, R. C. Concretization and sentence learning. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Anderson, T. & Armbruster, B. *Reader and text-studying strategies.* Conference on Reading Expository Materials, Research and Development Center, University of Wisconsin, Madison, Wisconsin, November 11-12, 1980.

Anderson, R. C. & Biddle, W. B. On asking people questions about what they are reading. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Anderson, I. H. & Dearborn, W. F. *The psychology of teaching reading.* New York: Ronald, 1952.

Anderson, R. C. & Pichert, J. W. *Recall of previously unrecallable information following a shift of perspective.* (Tech. Rep. No. 41). Champaign-Urbana, Illinois: Center for the Study of Reading, University of Illinois, April, 1977.

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior,* 1976, *15,* 667-679.

Anderson, R. C., Spiro, R. J., & Montague, W. E. *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Erlbaum, 1977.

Andre, M. E., & Anderson, T. H. The development and evaluation of a self-questioning study technique. *Reading Research Quarterly,* 1978-79, *14,* 605-623.

Athey, I. Reading research in the affective domain. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Ausubel, D. P. Use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology,* 1960, *51,* 267-272.

Bartlett, F. C. *Remembering.* Cambridge, England: Cambridge University Press, 1932.

Bond, E. *Reading and ninth grade achievement.* New York: Bureau of Publications, Teachers College, Columbia University, 1938.

Bond, G. L. & Dykstra, R. The cooperative research program in first-grade reading instruction. *Reading Research Quarterly,* 1967, *2,* 5-142.

Bransford, J. C. & Franks, J. J. Memory for syntactic form as a function of semantic context. *Journal of Experimental Psychology,* 1974, *103,* 1037-1039.

Bransford, J. D. & Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior,* 1972, *11,* 717-726.

Broudy, H. S. Types of knowledge and purposes of education. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, New Jersey: Erlbaum, 1977.

Brown, A. L. Development, schooling, and the acquisition of knowledge about knowledge: Comments on Chapter 7 by Nelson. In R. C. Anderson, R. J. Spiro, W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Erlbaum, 1977.

Buswell, G. T. *Fundamental reading habits: A study of their development.* Supplementary Educational Monographs, No. 21. Chicago: University of Chicago Press, 1922.

Buswell, G. T. Lecture in course on the Improvement of Reading, University of California, Berkeley, 1952.

Campbell, D. T. & Stanley, J. C. *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally, 1963.

Carroll, J. B. A model of school learning. *Teachers College Record,* 1963, *64,* 723-733.

Cattell, J. McK. The time it takes to see and name objects. *Mind,* 1886, *11,* 63-65.

Chomsky, N. *Syntactic structures*. The Hague: Mouton, 1957.

Chomsky, N. Review of B. K. Skinner's, *Verbal behavior. Language*, 1959, *35*, 26-58.

Clark, H. H. & Clark, E. V. *Psychology and language*. San Francisco: Harcourt, Brace & Jovanovich, 1977.

deCharms, R. Personal causation training in the schools. *Journal of Applied Social Psychology*, 1972, *2*, 95-113.

Doctorow, M., Wittrock, M. C., and Marks, C. Generative processes in reading comprehension. *Journal of Educational Psychology*, 1978, *70*, 118-119.

Donlan, D. & Singer, H. Active comprehension of short stories. In M. P. Douglass (Ed.), *Forty-Third Yearbook of the Claremont Reading Conference*. Claremont, California: The Claremont Reading Conference, Claremont Graduate School, 1979.

Dreher, M. J. *The instantiation of general terms in children*. Unpublished doctoral dissertation, University of California, Riverside, 1980.

Earle, R. & Barron, R. F. An approach for teaching vocabulary in content subjects. In H. L. Herber & R. F. Barron (Eds.), *Research in reading in the content areas: Second year report*. Syracuse, New York: Reading and Language Arts Center, Syracuse University, 1973.

Frase, L. T. Learning from prose material: Length of passage, knowledge of results, and position of questions. *Journal of Educational Psychology*, 1967, *58*, 266-272.

Frase, L. T. Boundary conditions for mathemagenic behaviors. *Review of Educational Research*, 1970, *40*, 337-348.

Frase, L. T. & Schwartz, B. J. Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 1975, *67*, 628-635.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458.

Freedle, R. O. *New directions in discourse Processing*, Volume II in the Series, *Advances in discourse Processing*. Norwood, N.J.: Ablex, 1979.

Goodman, K. S. A psycholinguistic view of reading comprehension. In G. B. Schick and M. M. May (Eds.), *New frontiers in college-adult reading*. Fifteenth Yearbook of the National Reading Conference. Marquette, Wisconsin: The National Reading Conference, 1966.

Goodman, K. S. Behind the eye: What happens in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*, Second Edition. Newark, Delaware: International Reading Association, 1976.

Gough, P. One second of reading. In H. Singer and R. B. Ruddell (Eds.). *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Hansen, J. & Pearson, D. *The effects of inference training and practice on young children's comprehension*. (Tech. Rep. No. 166). Champaign, Illinois: Center for the Study of Reading, University of Illinois, 1980.

Hayes, D. A. *The effect of text-embedded analogy upon comprehension and learning*. Unpublished doctoral dissertation, University of Arizona, 1979.

Helfeldt, J. P. & Lalik, R. Reciprocal student-teacher questioning. *Reading Teacher*, 1976, *30*, 283-287.

Herber, H. L. *Teaching reading in content areas*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Holmes, J. A. *Factors underlying major reading disabilities at the college level*. Doctoral dissertation, University of California, Berkeley, 1948. Published under the same title in *Genetic Psychology Monographs*, 1954, *49*, 3-95.

Holmes, J. A. *The substrata-factor theory of reading*. Berkeley, California: California Book Co., 1953. (out of print)

Huey, E. B. *The psychology and pedagogy of reading*. New York: Macmillan, 1908.

Javal, E. Essai sur la physiologie de lecture. *Annales d' Oculistique*, 1879, *82*, 242-253.

Judd, C. H. & Buswell, G. T. *Silent reading: A study of the various types. Supplementary Educational Monographs, No. 23*. Chicago: University of Chicago Press, 1922.

Katz, I. C. *The effects of instructional methods on reading acquisition systems*. Unpublished doctoral dissertation, University of California, Riverside, 1980.

Kavanagh, J. F. & Mattingly, I. G. *Language by ear and by eye*. Cambridge, Mass.: MIT Press, 1972.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum-Wiley, 1974.

Kintsch, W. *Text representations*. Conference on Reading Expository Materials, Research, and Development Center, University of Wisconsin, Madison, Wisconsin, November 11-12, 1980.

LaBerge, D. & Samuels, S. J. Toward a theory of automatic information processing in reading. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: The International Reading Association, 1976.

LaBerge, D. & Samuels, S. J. *Basic processes in reading: Perception and comprehension*. Hillsdale, N.J.: Erlbaum, 1977.

Levin, J. R. Determining sample size for planned and post hoc analysis of variance comparisons. *Journal of Educational Measurement*, 1975, *12*, 99-107.

Levin, J. R. Comprehending what we read: An outsider looks in. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Levin, J. R. *Do pictures improve children's prose learning? An examination of the evidence.* Invitational paper presented at the Twenty-Seventh Annual Meeting of the National Reading Conference, New Orleans, Louisiana, December 2, 1977.

Levin, H. & Gibson, E. *The psychology of reading*. Cambridge, Mass.: MIT Press, 1975.

Levin, H. & Williams, J. P. *Basic studies on reading*. New York: Basic Books, 1970.

McConkie, G. W. & Rayner, K. Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Manzo, A. V. *Improving reading comprehension through reciprocal questioning*. (Doctoral dissertation, Syracuse University, 1970). *Dissertation Abstracts International*, 1970, *30*, 5344A (University Microfilms No. 70-10, 364).

Marks, C. B., Doctorow, M. J., & Wittrock, M. C. Word frequency and reading comprehension. *Journal of Educational Research*, 1974, *67*, 259-262.

Marshall, N. *The structure of semantic memory for text*. Unpublished doctoral dissertation, Cornell University, Ithaca, New York, 1975.

Meyer, B. *The organization of prose and its effect upon memory*. Amsterdam, The Netherlands: North Holland, 1975.

Meyer, B., Brandt, D., & Bluth, G. Use of top-level structure in text: Key for reading comprehension of ninth grade students. *Reading Research Quarterly*, 1980, *16*, 72-103.

Miyake, N. & Norman, D. A. To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, 1979, *18*, 357-364.

Nicholson, T. *The relative effects of different error types on understanding of connected discourse.* Unpublished doctoral dissertation, University of Minnesota, 1977.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Pichert, J. & Anderson, R. C. Taking different perspectives on a story. *Journal of Educational Psychology*, 1977, *69*, 309-315.

Robinson, F. P. *Effective study*. New York: Harper & Brothers, 1946.

Robinson, F. P. Study skills for superior students in secondary school. *The Reading Teacher*, 1961, *15*, 29-33.

Robinson, H. A. *Teaching reading and study strategies: The content areas*. Boston: Mass.: Allyn and Bacon, 1975.

Robinson, F. P. & Hall, P. Studies of higher-level reading abilities. *Journal of Educational Psychology*, 1941, *32*, 445-451.

Rogers-Zegarra, N. *Anglo and Chicano literal and scriptal comprehension of ethnic stories*. Paper presented at the National Reading Conference, San Diego, California, December 3-6, 1980.

Rosenthal, T. L., Zimmerman, B. J., & Durning, K. Observationally induced changes in children's interrogative classes. *Journal of Personality and Social Psychology*, 1970, *16*, 681-688.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. *American Educational Research Journal*, 1966, *3*, 241-249.

Rothkopf, E. Z. *Adjunct aids, and the control of mathemagenic activities during purposeful reading.* Conference on Reading Expository Materials, Research and Development Center, University of Wisconsin, Madison, Wisconsin, November 11-12, 1980.

Rothkopf, E. Z. & Kaplan, R. An exploration of the effective density and specificity of instructional objectives on learning from text. *Journal of Educational Psychology*, 1972, 295-302.

Royer, J. M. & Cable, G. W. Facilitated learning in connected discourse. *Journal of Educational Psychology*, 1975, *67*, 116-123.

Ruddell, R. B. The effect of similarity of oral and written patterns of language structure on reading comprehension. *Elementary English*, 1965, *42*, 845-851.

Ruddell, R. B. Language acquisition and the reading process. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware. International Reading Association, 1976.

Rumelhart, D. E. Notes on a schema for stories. In D. Bobrow & A. Collins (Eds.), *Representation and* *understanding: Studies in cognitive science.* New York: Academic Press, 1975.

Rumelhart, D. E. *Toward an interactive model of reading.* (Tech. Rep. No. 56). San Diego: Center for Human Information Processing, University of California, 1976.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels, *Basic processes in reading: Perception and comprehension.* Hillsdale, N.J.: Erlbaum, 1977.

Rumelhart, D. *Processes of comprehending brief stories.* Invitational paper presented at California State College, San Bernardino, February 22, 1978.

Russell, D. H. *Children's thinking.* New York: Ginn, 1956.

Schank, R. C. & Abelson, R. P. *Scripts, plans, goals, and understandings.* Hillsdale, N.J.: Erlbaum, 1977.

Singer, H. Substrata-factor theory of reading: Theoretical design for teaching reading. In H. Singer and R. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Singer, H. Active comprehension: From answering to asking questions. *The Reading Teacher,* 1978, *31,* 901-908.

Singer, H. Active comprehension: From answering to asking questions. Reprinted in C. McCullough (Ed.), *Inchworm, inchworm: Persistent questions in reading education.* Newark, Delaware: International Reading Association, 1979.

Singer, H. Discussion of Walter Kintsch's *Text Representations.* Conference on Reading Expository Materials, Research and Development Center, University of Wisconsin, Madison, Wisconsin, November 11-12, 1980(a).

Singer, H. Towards an instructional theory for learning from text: A discussion of Ernst Rothkopf's *Adjunct aids, and the control of mathemagenic activities during reading.* Conference on Reading Expository Materials, Research and Development Center, University of Wisconsin, Madison, Wisconsin, November 11-12, 1980(b).

Singer, H. & Donlan, D. *Reading and learning from text.* Boston: Little, Brown, 1980.

Singer, H. & Ruddell, R. B. (Eds.) *Theoretical models and processes of reading.* Newark, Delaware: International Reading Association, 1970. Second edition, 1976.

Smith, N. B. Patterns of writing in different subject areas, Part I. *Journal of Reading,* 1964, *8,* 31-37. (a)

Smith, N. B. Patterns of writing in different subject areas, Part II. *Journal of Reading,* 1964, *8,* 92-102. (b)

Strang, R. *Exploration in reading patterns.* Chicago: University of Chicago Press, 1942.

Strange, M. C. The effect of orthographic anomalies upon good reading behavior. *Journal of Reading Behavior,* 1979, *11,* 153-161.

Taba, H. The teaching of thinking. *Elementary English,* 1965, *42,* 534-542.

Taylor. B. M. Good and poor readers' recall of familiar and unfamiliar text. *Journal of Reading Behavior,* 1979, *11,* 375-380.

Thorndike, E. Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology,* 1917, *8,* 323–332.

Thurstone, L. Note on a reanalysis of Davis's Reading Test. *Psychometrika,* 1946, *2,* 185-188.

Tierney, R. J. & LaZansky, J. The rights and responsibilities of readers and writers: A contractual agreement. *Language Arts,* 1980, *57,* 606-613.

Tierney, R. J. & Spiro, R. J. Some basic notions about reading comprehension: Implications for teachers. In J. Harste & R. F. Carey (Eds.), *New perspectives on comprehension,* Monograph No. 3, Language and Reading Series. Bloomington: Indiana University, 1979.

Tzeng, C., & Singer, H. *Perception of print.* Hillsdale, N. J.: Erlbaum, 1981.

Van Dijk, & Kintsch, W. Cognitive psychology and discourse: recalling and summarizing stories. In W. A. Dressler (Ed.), *Current trends in textlinguistics.* Berlin, New York: deGruyter, 1977.

Wanat, S. F. Relations between language and visual processing. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Winograd, T. Understanding natural language. *Cognitive Psychology,* 1972, *3,* 1-191.

Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* New York: Henry Holt, 1954.

ANN L. BROWN
University of Illinois

METACOGNITION: THE DEVELOPMENT OF SELECTIVE ATTENTION STRATEGIES FOR LEARNING FROM TEXTS

I. INTRODUCTION

In this paper, I will discuss some recent research from my laboratory concerned with strategies for studying. In so doing I will address various forms of behavior that could be called metacognitive (Baker & Brown, in press a) but needn't be. My primary aim is to demonstrate the complex interaction of many forms of knowledge that take place when students are engaged in a variety of learning activities, including learning from reading.

A. Metacognition

First, however, what is metacognition and how does it relate to reading? In recent reviews of the literature, Linda Baker and I (Baker & Brown, in press a, b) argued that it is necessary to distinguish between two types of metacognitive knowledge: 1) knowledge about cognition and 2) regulation of cognition. These two forms of metacognition are closely related. Attempts to separate them lead to oversimplification but such a separation does help us clarify the topic of discussion.

The first cluster is knowledge about cognition. Here the focus is on the knowledge readers have about their own cognitive resources and the compatibility between themselves as readers and the demands of a variety of reading situations. This form of self-knowledge is relatively stable, statable, fallible and late-developing. It is the form of knowledge that has traditionally been referred to as knowing that. It is *stable* because one would expect a reader who knows pertinent facts (such as that passages containing familiar words are easier to read than those containing unfamiliar words) to continue to "know" these facts when interrogated appropriately. This form of knowledge is often *statable* in that the child can reflect on the cognitive processes involved in reading and discuss them with others. Of course, this form of knowledge is also *fallible* in that the child, or adult for that matter, can perfectly well "know" certain facts about reading that are not true!

The ability to reflect on one's own cognitive processes, to be aware of one's own activities while reading, writing, solving problems, etc. is a *late-developing* skill with important implications for the child's effectiveness as an active, planful learner. If children are aware of what is needed to perform effectively, then it is possible for them to take steps to meet the demands of a learning situation more adequately. If, however, children are not aware of their own limitations as learners or the complexity of the task at hand, then they can hardly be expected to take preventive actions in order to anticipate or recover from problems.

The second cluster of metacognitive activities, regulation of cognition, consists of processes that are relatively unstable, rarely statable, without considerable effort, and relatively age independent. It consists of the self-regulatory mechanisms used by an active learner during an ongoing attempt to solve problems. These indices of metacognition include *planning* one's next move,

32

checking the outcome of any strategies one might use, *monitoring* the effectiveness of any attempted action, *testing, revising* and *evaluating* one's strategies for learning. These are not necessarily stable skills in the sense that although they are more often used by older children and adults, they are not always used by them, and quite young children may monitor their activities on a simple problem. Indeed, a case could be made that all learning involves monitoring to some degree. They are also not often statable, we know how to do a great many things that we cannot state explicitly; this is the form of knowledge often referred to as knowing how.

Expert readers monitor their comprehension and retention and evaluate their own progress in the light of the purposes for which they are reading. With repeated experience on these routine school activities, many of these cognitive monitoring processes become automatic. That is, although mature readers typically engage in comprehension monitoring, it is not often or even usually a conscious experience (Brown, 1980, in press a). When comprehension is proceeding smoothly, good readers proceed as if on automatic pilot, until a problem is detected. Some triggering event (Brown, 1980; Collins, Brown, Morgan & Brewer, Note 1) alerts them to a comprehension failure. Then and only then does the understanding process slow down and become planful, demanding conscious effort. Study monitoring for the expert involves many automatic, overlearned components, although here the need for conscious forethought, planning, and checking is somewhat greater. Less successful students are much less aware of the need to be strategic, plan ahead, monitor and check their own understanding. They have not yet learned how to learn from texts (Brown, in press a). Reading is not a primary or preferred mode of obtaining information and the task of studying is often interpreted to involve nothing more than passive, sometimes desperate, rereading. The self-regulatory activities that they can perfectly well apply to listening, for example, they are not yet able to apply to reading.

The importance of self-awareness and self-control while learning were recognized well before the current popularity of metacognitive theories; indeed, educational psychologists since the turn of the century (e.g., Dewey, 1910; Huey, 1968; Thorndike, 1917) were quite aware that reading involves the planning, checking, and evaluating activities now called metacognitive skills. For example, Dewey's system of inducing reflective thinking is essentially a call for metacognitive training. The aim was to induce active monitoring, critical evaluation and deliberate "seeking after meaning and relationships." Similarly, Thorndike claimed that reading was reasoning and involved a great many activities now called metacognitive.

> Understanding a paragraph is like solving a math problem. It consists of selecting the right elements of the situation and putting them together in the right relations, and also with the right amount of weight or influence or force for each. The mind is assailed as it were by every word in the paragraph. It must select, repress, soften, emphasize, correlate and organize, all under the influence of the right mental set or purpose or demand. (Thorndike, 1917, p. 329)

Although the term metacognitive may be new, the type of knowledge to which it refers has long been recognized.

B. Study Strategies and Selective Attention

Also not new is the main focus of the work I will concentrate on in this paper, study strategies. There is a long history of interest in the types of knowledge and strategies students bring to the task of learning from texts—note-taking, underlining, adjunct aids, question-asking, outlining, etc. (Anderson & Armbruster, 1980; Brown, in press a). In the past, however, the work has been limited in several ways: first, the majority of prior work has concentrated almost exclusively on the study activity of adults, usually college students. Second, the majority of studies have been descriptive rather than explanatory or manipulative. And third, the most important limitation is that the majority of studies have shared a concentration on product rather than process. That is, the main metric has been some outcome measure such as the test scores of students who do or do not use a certain strategy. Little or no consideration is given to the activity of the studiers, i.e., what they are *actually doing* while studying. An example of this is the large number of experiments that have considered the product of the activity of note-takers, i.e., how well they do on tests, but have ignored the processes that the note-taker employs; even the notes themselves are not examined for evidence of the process that the taker of notes might be employing (Brown & Smiley, 1978).

Developmental psychologists have been concerned with the processes that learners employ when studying. However, their work also has its limitations; primarily the fact that the main concentration is on children learning strange kinds of material in even stranger settings. It is only recently that developmental psychologists have turned their attention to the school-like activity of studying texts (Brown & Smiley, 1978; Brown & Day, Note 2). One impetus for the work carried out in my laboratory has been an attempt to combine the interest in developmental processes of learning with the concentration of educational psychologists on learning in academic settings.

To be very brief, the major findings from the developmental literature that have influenced our work are: first, the notion of *Selective Attention*. As children mature they become adept at channeling their attention to the most informative aspect of information. They become able to ignore irrelevant or less informative aspects of material. This selective attention is influenced by three main factors: 1) *Strategies*, as children mature they show a marked tendency to introduce a variety of increasingly ingenious ploys to aid learning. 2) *Metacognition*, as previously discussed, as children mature they acquire more knowledge concerning methods and procedures for studying and the match between the task at hand and their available repertoire of strategies, and 3) *Content Knowledge*, obviously as children mature, they acquire more knowledge about the world around them. This increased content knowledge must influence how they selectively attend, and how they tailor their learning process.

C. *Plan of the Chapter*

The work from my laboratory has concentrated on examining the mutual compatibility and interdependence of these forms of knowledge and the effect they have on students' efficiency when learning from texts. In the remaining section of the paper I will give a brief description of four lines of work: 1) Selecting and studying the main idea by means of such strategies as note-taking and underlining; 2) filling in the details to support the main idea; 3) summary writing and 4) question asking. On the last two topics I will discuss initial attempts to train students to become more effective users of the necessary strategies.

II. SELECTING AND STUDYING THE MAIN IDEAS

The original work on text learning was conducted with Sandra Smiley. We began in 1972. At that time there was little existing developmental work on children's memory for stories, with the exception of the classic studies by Binet (Binet & Henri, 1894) and Piaget (1928). Influenced by Bartlett's work with adults, we chose folk-tales as our story material and for the first half of the paper I will discuss our work on children learning story materials. In the second half of the paper my concern will be children studying expository materials.

Students are commonly exhorted to concentrate on the main ideas when studying, but in order to be responsive to this suggestion, they must be aware of what the main points of a passage are. This is a gradually developing skill; and although children as young as six can often indicate the main character and sequence of events in a simple narrative, they often experience difficulty isolating central issues in more complex prose (Brown & Smiley, 1977; Smiley, Oakley, Worthen, Campione, & Brown, 1977). For example, we found that grade school children who were perfectly able to recall the main ideas of folk stories had much more difficulty rating sections of the stories in terms of their importance to the theme. We asked students, aged 8, 10, 12, and 18 years, to rate the idea units of complex folk stories in terms of their importance to the theme of a passage (Brown & Smiley, 1977). Only 18-year-olds could reliably distinguish the previously rated four levels of importance. Twelve-year-olds did not differentiate the two intermediate levels of importance, but they did assign their lowest scores to the least important and highest scores to the most important elements. Eight-year-olds made no reliable distinction between levels of importance in their ratings and even 10-year-old students could only distinguish the highest level of importance from all other levels. There was considerable agreement between independent groups of college students, and even 12-year-olds, concerning the importance of constituent idea units of a text passage; but 8- and 10-year-old subjects were unable to differentiate units in terms of their relative importance to the text. Thus, when the passage is relatively long (55 idea units) and complex (fifth-grade reading level), even junior high school students experience some difficulty in identifying the important elements of text.

It is important to note, however, that recall of the passages at all ages was sensitive to the importance level of the units. After rating a story for the importance of its units, subjects were asked to read and recall another story. Although the older students remembered more than the younger children, the general pattern of results was consistent across the age range of 8 to 18 years, the least important units were recalled less frequently than all other units and the most important units were most often recalled. Even without prior awareness of the importance of constituent units, younger children still favor the important units in recall. Apparently we spontaneously abstract the main ideas of an oral or written communication even when no deliberate attempt to do so is instigated.

These initial findings have important implications for studying. In order to go beyond retention of just a few main points, i.e., to achieve a more complete "fleshed-out" memory of the text, one must engage in active strategies to ensure increased attention to important material that will not be retained automatically. This need for active intervention is particularly pressing if the contents of lengthy texts are to be retained over some period of time, a typical school learning situation. If young children have difficulty distinguishing what is important they will also have difficulty studying. Quite simply, one cannot selectively attend to importance in the absence of a fine sensitivity to what *is* important.

As children mature they become better able to identify what are the essential organizing features and crucial elements of texts. Thanks to this foreknowledge, they make better use of extended study time. For example, we found that when given an extra period for study (equal to three times their reading rate), children from seventh grade up improved their recall considerably for important elements of text; recall of less important details did not improve. Children below seventh grade did not usually show such effective use of additional study time; their recall improved, if at all, evenly across all levels of importance. As a result, older students' recall protocols following study included all the essential elements and little trivia. Younger children's recall, though still favoring important elements, had many such elements missing (Brown & Smiley, 1978).

We believe that the older students benefit from increased study time as a direct result of their insights into their own cognitive processes while reading and their ability to predict ahead of time what are the important elements of the texts. Younger students, not so prescient, cannot be expected to distribute extra time intelligently; they do not concentrate on only the important elements of text, since they do not know in advance what they are.

To substantiate our belief that metacognitive control governs this developmental trend we observed the study actions of the students. In particular, we examined the physical records they provided—notes and underlining of texts. A certain proportion of children from fifth grade and up spontaneously underlined or took notes during study. At all ages, the physical records of spontaneous subjects favored the important elements; i.e., the notes or underlined sections concentrated on elements of the text previously rated as crucial to the theme (Brown & Smiley, 1978).

Students induced to adopt one of these strategies did not show a similar sensitivity to importance; they took notes or underlined more randomly. Although the efficiency of physical record keeping in induced subjects did improve with age, it never reached the standard set by spontaneous users of the strategy. Furthermore, the recall of spontaneous producers of the strategies was much superior. Even fifth graders who spontaneously underlined showed an adult-like pattern and used extra study to differentially improve their recall of important elements. This interesting difference between spontaneous and induced users of the note-taking or underlining strategy may explain the curious lack of effect so far found in adult studies concerned with these processes (Anderson, 1980). Previous studies have typically assigned students arbitrarily to treatment groups, thus, in effect, combining those who would or would not use the strategy on their own volition. This procedure tends to mask the efficiency of the strategy, an efficiency which is dependent on its subject-generated nature (Brown, 1980, in press a; Brown & Smiley, 1978).

III. FILLING IN DETAILS

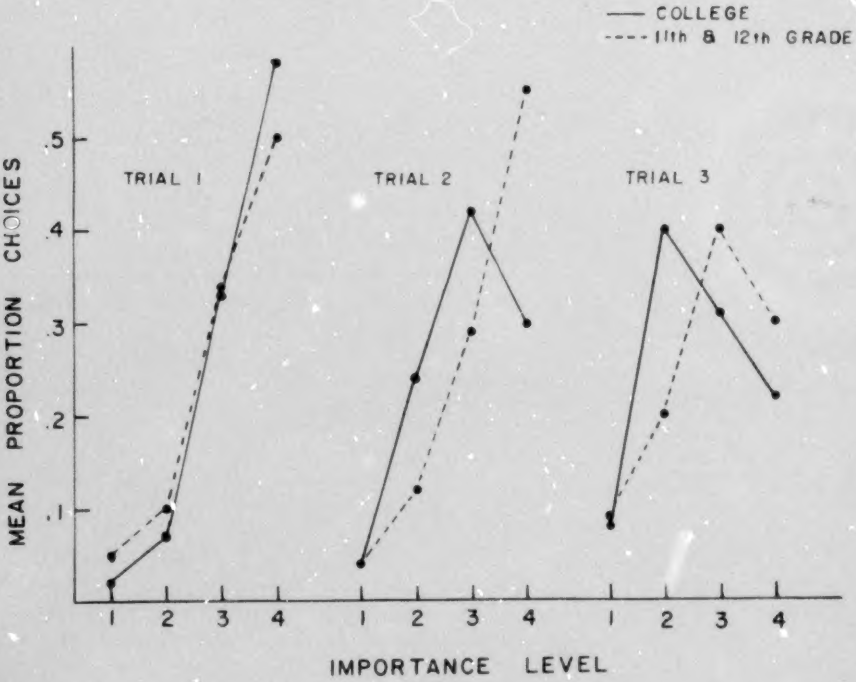
Once they have ensured that the main ideas are well understood, efficient students will take steps to fill in the details. One method of achieving this is to test oneself to determine which details one has failed to recall and then to devote extra attention to the previously missed information. Even grade school children can initiate some efforts to test their current state of memory (Masur, McIntyre, &

Flavell, 1973) and mildly retarded children can be taught to do so (Brown & Campione, 1977; Brown, Campione, & Barclay, 1979). But this can become a much more difficult task if the material is complex.

We examined the ability to concentrate on information one has previously failed to recall, in a laboratory task analogous to the process of studying (Brown, Smiley, & Lawton, 1977; Brown & Campione, 1979). Students from fifth through twelfth grade, together with college students, were asked to study prose passages until they could recall all the details in their own words. They were allowed repeated study trials. The passages were divided into constituent idea units rated in terms of their importance to the theme; there were four levels of rated importance. On each trial the students were allowed to select a subset of the idea units (printed on cards) to keep with them while they attempted recall. After recall and a rest period, the entire process was repeated.

Of interest are the idea units selected as retrieval aids to enhance recall. The data for college students and older high school students are shown in Figure 1. On the first trial the majority of students at all ages selected the most important units to help them recall. Children below high school age continued to do this, even though across trials they became perfectly able to recall the most important information without aid, but persistently failed to recall additional details. College students, however, modified their selection as a function of trials: on the first trial

FIGURE 1 Retrieval Cue Selection over Trials for the Older Students (from Brown & Campione, 1979)



they selected predominantly important Level 4 units for retrieval aids. On the second trial they shifted to a preference for Level 3 units while on the third trial they preferred Level 2 units. On all three trials, Level 1 units were treated appropriately as trivia. As they learned more and more of the material, college students shifted their choice of retrieval cues to reflect their estimated state of knowledge.

Older high school students showed the same basic pattern as the college students, but they were one trial behind; they did not begin to shift to less important units until the third trial. This lag could be due to slower learning, i.e., both groups shifted when they reached the same criterion of learning but the younger students took an extra trial to reach that criterion. It could also be due to a slower selection of the effective study strategy of switching to less important units, i.e., both groups learned as much on each trial but it took high school students longer to realize they needed to shift cue selection. The latter appears to be the correct interpretation, for even when students were matched on the basis of degree of learning, the younger students still took longer to shift their choice of retrieval cues.

The ability to select suitable retrieval cues is a late-developing skill, because it requires a fine degree of sensitivity to the demands of studying. The successful user of the flexible retrieval plan illustrated in these studies must have (a) information concerning the current state of knowledge, i.e., what he knows of the text and what he does not yet know; (b) knowledge of the fine gradation of importance of various elements of texts, i.e., what is important to know and what can be disregarded; and (c) the strategic knowledge to select for retrieval cue information previously missed.

It is interesting to note that in some sense the very late development of this study skill is surprising. Students as young as third grade (Masur *et al.*, 1973) and even fifth grade educably retarded students (Brown & Campione, 1977) can employ a primitive version of this strategy. Given a list of pictures to remember, they do select for extra study items they have failed to recall. But it is not until much later in their school careers that students can apply this strategy to the complex task of learning from texts. I believe that this is a ubiquitous problem for young studiers who must learn to apply strategies they can use perfectly well outside of reading tasks to the less familiar written medium. We are currently working on training students to use the appropriate selection strategy for filling in details by illustrating the activity on the simple list learning task and explicitly stating the obvious analogy to the text learning situation (Brown, Jones, & Mancuso, work in progress).

IV. SUMMARY WRITING

Jeanne Day and I have been working on a series of studies concerning the development of the ability to write summaries of studies and expository materials (Brown, Campione, & Day, 1981; Brown & Day, Note 2; Day, 1980). As we have seen, adequate recall of anything more than simple short stories requires both effort and judgment. Recall, if it is to include more than a bare skeleton, demands strategies for concentrating on difficult and important elements (Brown & Campione, 1979); it requires judgment of what to include and what to omit. For these reasons, improved recall after studying complex texts is a relatively late developing skill (Brown & Smiley, 1978). Similarly, summarizing texts must also entail judgment and effort if more than the barest synopsis is required. Estimations of fine degrees of relative importance must be made and rules for condensation employed. Thus, we predicted that the ability to provide an adequate summary of a lengthy text would also be a late developing skill, a prediction well supported by the existing traditional educational literature (Germane, 1921a, 1921b; Stordahl & Christensen, 1956).

A problem with estimating children's ability to summarize is that it is essential that we are able to distinguish between the "automatic" by-product of comprehension, and the deliberate result of judgment and effort (Brown, 1975; Hasher & Zacks, 1979). If children as well as adults "automatically" extract the main gist when comprehending stories, there is a danger that we will obtain excellent summaries from children, not because they thought about their responses, but because they produced all the information they had available. It is necessary, therefore, to ensure that subjects can recall much of the information they are required to summarize. Two ways of accomplishing this are: 1) to use simple stories which are well-formed according to a story grammar view and that produce excellent recall (Johnson, 1978). But such well-formed stories are already excellent summaries, leaving little room for judgment and selection. The second alternative, the one we adopted, is to use more lengthy, complex stories and require that the students learn the texts to some criterion before preparing a summary. Under these circumstances, we can examine the development of children's judgments concerning what elements to omit or include but we would not have confounded memory and selection.

Six studies will be reported briefly. The first serves as a descriptive study of the development of the ability to summarize texts. Here we build upon our previous work on children's identification and use of important element of texts (Brown & Campione, 1979; Brown & Smiley, 1977, 1978; Brown, Smiley, & Lawton, 1977; Brown, Smiley, Day, Townsend, & Lawton, 1977). The second study is a more analytic examination of the condensation rules used by children and adults to achieve adequate synopses of texts. In the third study we examined experts' use of summarization rules using on-line "talk-aloud" protocols. Following our consideration of experts we turned our attention to novices; in the fourth study we examined the potential diagnostic power of our developmental norms by considering the performance of junior college students, a population known to experience problems in critical reading and effective studying. In the fifth and sixth experiments we attempted to train a variety of junior college students to improve their ability to summarize texts.

A. *Summarizing the Main Idea.* We asked students from fifth, seventh and eleventh grade, as well as college-age students, to learn a story so that they could remember all the information, even the details, in their own words. Only those students able to comply with this request (remember more than 80% of all four levels of importance in our stories) were allowed to continue in the study. After recalling the story, the students were then asked to write a summary in their own words. The number of words was not specified.

After they had finished the first summary (and a break), the children were told that the editor (person in charge of space) had cut their space because there was a very important story that must be covered. Their task now was to write the story again but in only forty words. To help them they were given large sheets of paper with forty spaces at the bottom where they were to write their summary. No explicit instructions concerning the use of the top half of the page (blank) were given. We thought that this might be a hint that the subjects should make a rough draft before attempting to fill in the spaces. In the final phase, the procedure was repeated except that the subject's space was cut back to twenty words.

A brief synopsis of the summarization performance of eleventh graders and college students is that they included in their summaries primarily level 3 and 4 (important) units. While some lower level units (1-2) were included in the free paraphrase, these dropped out as soon as space pressure was exerted. Even though the pattern of results was very similar for eleventh graders and college students, college students' greater efficiency was reflected in their ability to include more idea units in the same number of words.

The fifth grade pattern was somewhat different but intriguing in light of prior data from this age group. Remember that in our initial study (Brown & Smiley, 1977), fifth graders had been asked to rate the units of these stories for importance. The fifth graders' ratings differed from older children in that they were only able to distinguish level 4 units as more important to the theme than the remaining units. This difference in rating data has been found to influence the note-taking (Brown & Smiley, 1978), underlining (Brown & Smiley, 1978), and retrieval-cue selection (Brown & Campione, 1979; Brown, Smiley, & Lawton, 1978) of fifth grade subjects. Whereas older subjects show a clear pattern of underlining, notetaking, and cue selection, reflecting all levels of importance of idea units, fifth graders only show a preference in their notes, underscored sections or selections for level 4 units over all the other units. This pattern of results has been taken as evidence of a close link between knowledge factors and performance. As the fifth graders only recognize level 4 units as more important, they concentrate only on these units when working on these texts. Older students distribute their attention as a function of the finer degrees of importance that they are able to recognize.

This same pattern was found here with the unconstrained summary. When left free to paraphrase in as many words as they wished, fifth graders included many more level 4 units in their summaries than any other level but showed no preference for level 3 over levels 1 or 2. However, as space pressure was exerted in the constrained summaries, the pattern changed in interesting ways. Constrained to forty words, the fifth graders dropped level 1 units; further limited to twenty words the fifth graders dropped the level 2 units and thus ended up with a pattern indistinguishable from older subjects. This dropping of units as a function of their rated importance is the first evidence we have, in this extensive series of studies, that fifth graders are in any way sensitive to fine differences in importance between levels 1, 2 and 3 of these stories. When severely pressed for space even the younger children are sensitive to fine degrees of importance, a sensitivity that comparable age children do not show when studying the texts (Brown & Smiley, 1977, 1978) or rating them under a variety of conditions (Brown, Smiley, & Lawton, 1978; Brown & Campione, 1979). This is a nice illustration of the fact that estimating a child's knowledge or "awareness" is not a simple task; the degree to which a student will be judged "aware" depends on the indices used to measure that awareness (Baker & Brown, in press a; Brown, 1978; Brown, 1980). Only when severely pressed for space do fifth graders indicate that the units of these texts differ in terms of their importance to the theme of the story.

In all phases of the initial summarization study we provided scratch paper and blank spaces for any attempts to rough draft the summaries before filling in the spaces. Students varied in the degree to which they spontaneously prepared rough drafts. A prototype rough draft would consist of a preliminary written version of the passage, often with the number of words indicated, then a rewrite, and sometimes more than one rewrite. The proportion of students attempting a rough draft increased with age (.13, .37, .46 and .63 respectively for grades five, seven, eleven and college students).

Another interesting index of planning was our informal observation that the younger children not only failed to make a rough draft but also appeared to run out of space, i.e., the children complained that they had no spaces left into which to fit the end of the story. To test this observation we divided the text into two halves so that the halves contained approximately the same number of idea units, distributed evenly over the levels of importance. We then calculated how many units of the first and second half of the story were included in the summaries of children who did and did not make rough drafts (referred to as the Plan and No Plan groups). There was a dramatic difference between planning and not-planning subjects at the lower two ages. Those younger children who prepared a rough draft showed no effects of first half, second half in their summaries. Subjects not thinking ahead, however, favored the first half of the story. This confirms the anecdotal report that younger children ran out of space because they failed to leave room for the second half of the story. This was not true for older subjects or for younger subjects who made a rough draft. Younger children who failed to make a rough draft tended to run out of space before completing their summary. College students who do not produce a rough draft have sufficient control of their activity to enable them to produce an adequate representation of both halves of the story.

To summarize, by fifth grade, children are able to attempt a summary of lengthy texts but clear developmental trends are found. College students and older high school students outperform younger children in their propensity to plan ahead, in their sensitivity to fine gradations of importance in the text, and in their ability to condense more idea units into the same number of words. As predicted, under circumstances when a summary (or a recall protocol) is not just a measure of automatic retention, the ability to recursively work on the information to render it as succinctly as possible requires judgment and effort, knowledge and strategies (Baker & Brown, in press a, b; Brown & Smiley, 1978). As such, the ability to provide an adequate summary of a lengthy text is a late developing skill that continues to be refined throughout the school years (Brown & Day, Note 2).

B. *Condensation Rules for Summary Writing:* At this point my colleagues and I became dissatisfied with our total reliance on the rated importance measure as an indication of efficiency. Although the rated importance metric was quite successful as a measure of where attention was being directed, it did not provide sensitive information concerning the sophisticated strategies students used to direct attention. For example, the younger students in the Brown and Smiley (1978) note-taking study produced a pattern of note-taking that differed qualitatively from the older students; their notes were closely related to the surface structure of the text, both in the order of occurrence of the idea units and in the near verbatim selection of words actually contained in the text. Older subjects not only showed a greater tendency to take notes and to take notes that were extremely sensitive to rated importance, they also demonstrated a greater ability to paraphrase and to rearrange order into topic clusters.

A very similar pattern emerged in our initial summarization study (Brown & Day, Note 2). Younger children tended to operate on material mainly by deleting or retaining surface elements of the texts. Older students relied on paraphrase and condensation rules to combine and rearrange idea units, to depart from the words actually present in the text. It was by such means that they were able to condense more idea units into fewer words. It is these strategies that the importance level rating procedure did not measure. And it is these activities that we wished to capture by a more sensitive metric.

After considering the qualitative differences between the younger and older students' summaries we identified a number of general rules that we believed to be the ones used by the more mature. To test this we asked expert summarizers (college rhetoric and journalism teachers) what rules they used to summarize. Although they were remarkably uninformative about the existence of set rules, a consideration of their "talk-aloud" protocols while actually summarizing confirmed that the rules we had identified were, indeed, the primary ones used by experts. In addition, contemporaneous with our work, Kintsch and van Dijk (1978) described a series of macrorules for comprehension of texts that had a great deal of similarity to our summarization rules. Thus, from both a practical and theoretical viewpoint, we had reason to believe in the psychological reality of our rule coding system.

Two of the six rules we identified involve the deletion of unnecessary material. One should obviously *delete* material that is *trivial,* and even grade school children are quite adept at this if the content of the material is familiar. One should also *delete* material that is important but is also *redundant.* Kintsch and van Dijk's system also includes these two deletion rules. Two of the rules of summarization involve the *substitution of a superordinate* term or event for a list of items or actions. For example, if a text contains a list such as: cats, dogs, goldfish, gerbils and parrots, one can substitute the term pets. Similarly, one can substitute a superordinate action for a list of subcomponents of that action, i.e., John went to London, for: John left the house, John went to the train station, John bought a ticket, etc., etc. These rules are roughly comparable to Kintsch and van Dijk's (1978) generalization rules. The two remaining rules have to do with providing a summary of the main constituent unit of text, the paragraph. The first rule is—*select a topic sentence,* if any, for this is the author's summary of the paragraph. We referred to this as the *selection rule,* but it is similar to Kintsch and van Dijk's integration rule. The final rule is—if there is no topic sentence, *invent* your own. These operations are roughly equivalent to Kintsch and van Dijk's (1978) construction rule. We would like to point out that although we identified these basic rules as specific strategies actually used by students when engaged in the formal task of summarizing, they are also the rules used by the more mature students in the note-taking studies conducted earlier (Brown & Smiley, 1978). We agree with Kintsch and van Dijk that the *macrorules* of *deletion, superordination, selection* and *invention* are general rules underlying comprehension of texts.

These operations are used freely by experts when summarizing texts, but do less sophisticated readers realize that these basic rules can be applied? To map the developmental progression associated with the use of the basic rules we examined the ability of students from grades 5, 7 and 10 and various college students to use the rules while summarizing. We used specially constructed texts that enabled us to predict when each rule should be applied or at least would be applied by experts (college rhetoric teachers). Even the youngest children were able to use the two deletion rules with above 90% accuracy, showing that they understood the basic idea behind a summary—get rid of unnecessary material. On the more complex rules, however, developmental differences were apparent. Students became increasingly adept at using the topic sentence selection rules, with college students performing extremely well. However, the most difficult rule, invention, was rarely used by fifth graders, used on only a third of appropriate occasions by tenth graders and on only half of the occasions when it was appropriate, even by college students.

Although, within the Kintsch and van Dijk (1978) model, there is no reason to suspect that the five rules of summarization would be of differential difficulty, our developmental pattern was quite clear. One explanation for the differential difficulty of the five rules is that they demand differing degrees of cognitive manipulation and they depart to a greater or lesser extent from the already existing strategy favored by the younger participants. We refer to this as the *copy-delete* strategy. Fifth graders tend to treat the task of summarizing as one of deciding if to include or delete elements that actually occur in the surface structure of the original text. This is the same general strategy employed by fifth and seventh grade note-takers (Brown & Smiley, 1978) and even second grade summarizers attempting to orally summarize simple stories (Johnson, 1978). The strategy is as follows: a) read text elements sequentially; b) decide for each element on inclusion or deletion; c) if inclusion is the verdict, copy it more or less verbatim from the text.

This simple strategy is used consistently in a variety of text processing tasks by 11-12 year old students and it works relatively well in that it gets the job done, it does result in a product that is recognizably a summary, an outline or a set of notes. It is because the simple copy-delete strategy is so generally applicable and meets with some success that it is difficult to get students to attempt more complex rules. Bereiter and Scardamalia (in press) have found that the simple "knowledge-telling" strategy in writing, preferred by 11-12 year olds, is also resistant to change, again because it is relatively successful and generally applicable.

Consider the five rules of deletion, superordination and topic sentence manipulation in terms of how far they depart from the copy-delete strategy. Obviously the easy deletion rules map straight onto the existing strategy. Copy-delete works quite well for superordination with the exception that the students must add a superordinate in place of a list. But in order to use the topic sentence rules appropriately the students must abandon either the sequential unit by unit approach (selection) or both the sequential approach and the copy-delete principle (invention). To use the selection rule the students must have some realization of the unique status of the topic sentence within paragraphs and write their summaries in support of topic sentences they have selected. This would demand disrupting the sequentiality rule and giving unique status to topic sentences, probably by selecting them first to form the scaffolding of the summary. This is the rule used by experts (Brown and Day, Note 2). The copy-delete rule still applies in that one can copy the selected topic sentence straight from the text.

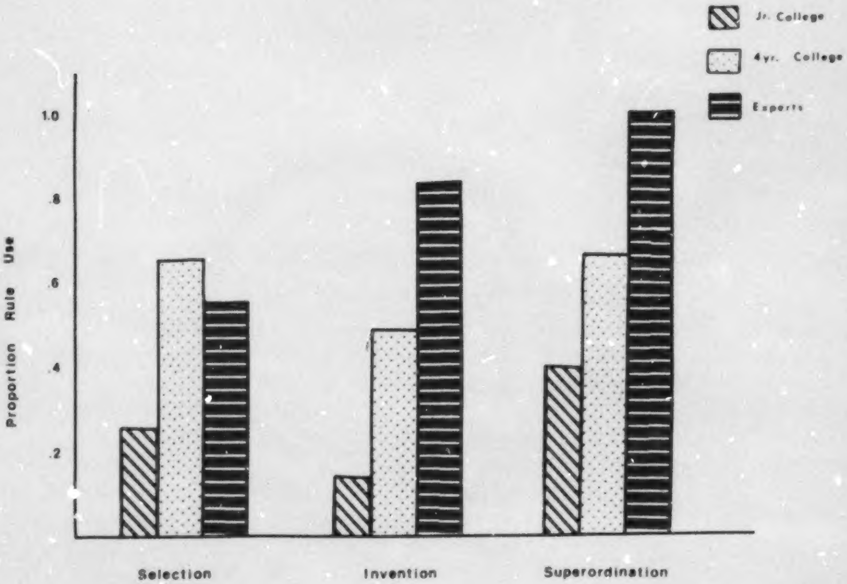
The difficult invention rule is, I think, difficult because it departs most radically from the favored copy-delete ploy. Students must now add something of their own, a synopsis in their own words of the implicit meaning of the paragraph. The invention rule, therefore, requires that the students add information rather than just delete, select or manipulate sentences already provided for them. It is these processes that are the essence of good summarization, that are used with facility by experts and that are most difficult for novice learners.

C. *Experts' Summaries:* Even college students did not apply the condensation rules perfectly, experiencing particular problems with the invention rule. We decided, therefore, to look at the performance of experts. Our experts were fourth-year graduate students in English who had taught rhetoric classes. Their performance on the summary task is presented in Figure 2, together with the data from undergraduates at a four-year college and junior college students.

The experts used the superordination rule perfectly compared to the 70% level set by college students. There were no differences between populations in the use of the selection rule. However, the experts used the difficult invention rule much more than did the college students (.84 vs. .49). Indeed, a case could be made that the experts performed perfectly because on the rare occasions that they did not receive a "correct score" for invention use, they had combined two paragraphs into one, thereby losing credit for one topic sentence use. (This tendency depressed their scores for the selection rule also.) Rarely did any of the other students combine paragraphs. Experts, however, favored the paragraph combining strategy and attempted to use it whenever possible.

We also collected verbal protocols from the experts. In the open-ended interviews prior to actually summarizing, the experts showed a surprising lack of evidence that they knew any effective rules for summarization. Their description of what a good summary was, and what to tell students, was essentially similar to

FIGURE 2 Use of the Selection, Invention and Superordination Rules by College Students with Varying Degrees of Expertise (from Brown & Day, Note 2)



that contained in rhetoric textbooks (Bessey & Coffin, 1934). They stressed that a summary is a concise statement of the theme and that one should avoid unnecessary repetition, be concise, include only main ideas, etc., but there was no mention of a systematic set of rules for doing this.

During their attempt to summarize, however, the experts made frequent mention of the basic rules. Of the statements judged to be a reference to rule use, 58% were an explicit statement of one of the five rules. For example, one expert provided the following succinct statement of the superordination rule: "One thing I've done is drop the kinds of plants. Instead of writing daisies, poppies, marigolds and lilies, all I've written is 'annual plants', again leaving out details and talking about generalizations." And for the invention rule, she provided: "The paragraph is about the cycle of the annual plants that produce seeds, waits until rainfall, bloom, produce seeds again, etc. Although it doesn't say so, all you need is to state this cycle, then you can drop the rest" (Brown & Day, Note 2).

The experts were unable or unwilling to give a precise statement of the rules prior to attempting to summarize a text; they spoke in very general terms about finding "main ideas" and "being concise", etc. However, when actually faced with the task of summarizing a passage, they were much more explicit about the rules they were employing and the rules they stated were, for the most part, the basic rules of deletion, superordination and topic sentence manipulation. The only other dominant rule that was used by experts and repeatedly appeared in their protocols was the combining-paragraphs rule. Experts used the rule routinely, whereas less mature students rarely attempted to combine across paragraphs, seeming instead to be "captured" by the paragraph structure provided in the original passage.

D. *Novices' Summaries:* Having examined experts' summarization performance, we turned our attention to novices. In order to examine the diagnostic value of our age norms, we repeated the summary study using junior college students, a population known for deficient skills of critical reading and studying, although a consideration of the traditional educational research literature would suggest that junior college students are not alone in their inability to adequately abbreviate text. Elementary school children (Germane, 1921a, 1921b), Air Force recruits (Stordahl & Christensen, 1956), and Harvard undergraduates (Perry, 1959) all demonstrate poor summarizing skills. In fact, summarizing is just one of several study techniques that these students fail to employ well (Anderson & Armbruster, 1980; Brown, in press a). Educators complain that high school students (Dynes, 1932; Beauchamp, 1923; Germane, 1921b), recruits for the armed forces (Weinstein, 1978; Stordahl & Christensen, 1956) and even college undergraduates (Preston & Tufts, 1948; McClusky & Dolch, 1924) lack effective note-taking, outlining and summarizing skills. An examination of the validity of these traditional claims, using our sensitive diagnosis of rule use, seemed timely. Our novices were junior college students enrolled in an English course that fulfilled the freshman rhetoric requirement at that college and at many four-year universities. That is, students could receive credit for this course should they continue their education at a four-year institution. The students were *not*, therefore, diagnosed as having any reading or writing problems on the basis of tests administered on entry to the college. In general, they were in a college preparation stream. English was their first language.

The novices performed extremely well on the deletion rules but departed significantly from the norms set by college students on the superordination, selection and invention rules (see Figure 2). In general, junior college students performed on a par with seventh graders in our developmental study. These data confirm the claims of educational psychologists that students from less academically privileged backgrounds perform poorly on a variety of text-processing strategies, including summarization. Our data take us beyond this global claim by providing a more fine grained analysis of where the students are experiencing particular problems. We see that the ability to delete trivial or redundant material is relatively intact, at least with the very simple expository materials used in these studies. The strategies needed for adequate manipulation of topic sentence rules are, however, much more problematic for these students. Junior college students, even those with no diagnosed reading or writing problems, perform on a level comparable to that of seventh graders from regular junior high schools (Brown & Day, Note 2). Given this depressing level of performance, it appears that there is ample room for improvement.

E. *Training Summarization Skills:* In her Ph.D. thesis, Jeanne Day undertook to train junior college students to use the condensation rules more appropriately. This was a complex study from which I will select only the highlights (for full details see Day, 1980). There were two studies in the Day thesis; the first provided three days of training in small groups to junior college students of two levels of ability: "*Normal*" students with no reading or writing problems diagnosed and *Remedial Writer* students with diagnosed writing problems but no known reading problems. The second study provided intensive, one-on-one training to a more severely learning impaired group, *Remedial Readers,* who had both reading and writing problems.

Students were assigned to one of four instructional conditions that varied in how explicit the training was: 1) *Self-Management:* the students were given general encouragement to write a good summary, to capture the main ideas, to dispense with trivia and all unnecessary words—but they were not told rules for achieving this end. This type of instruction is quite typical of the routine classroom practice for "teaching" summarization (Bessey & Coffin, 1934). 2) *Rules:* the students were given explicit instructions and modelling in the use of the rules. For example, they were given various colored pencils and shown how to delete redundant information in red, delete trivial information in blue, write in superordinates for any lists, underline topic sentences if provided and write in a topic sentence if needed. 3) *Rules Plus Self-Management:* the students in the third group were given both the general self-management instruction of Group I and the rules instruction of Group II, but they were left to integrate the two sets of information for themselves. 4) *Control Plus Monitoring:* the fourth and most explicit training condition involved training in the rules and additional explicit training in the monitoring of these rules—i.e., the students were shown how to check that they had a topic sentence for each paragraph (either underlined or written in), to check that all redundancies were deleted, all trivia erased, etc. and that any lists of items were replaced with superordinates; they were taught how to monitor that the rules had been used correctly and exhaustively.

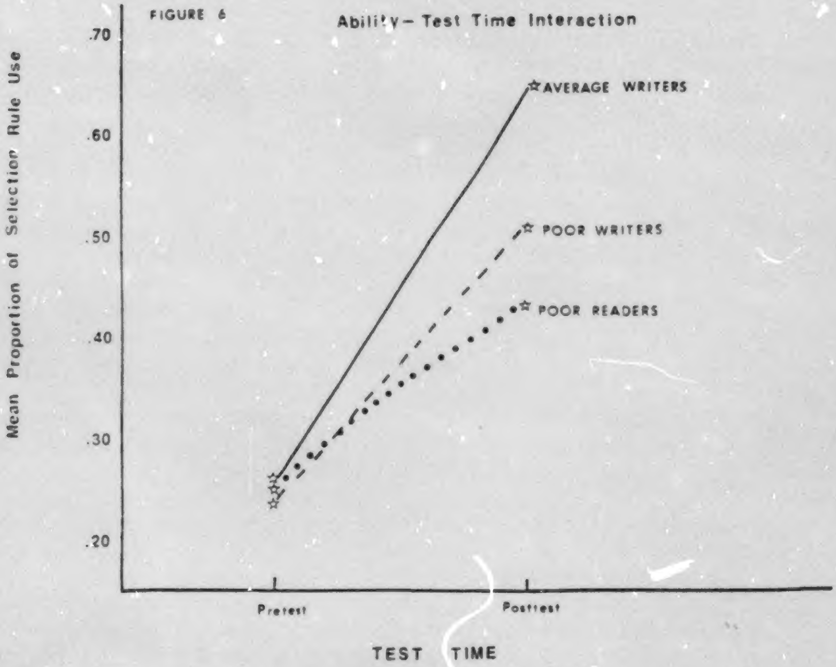
The results of training were encouraging. The two deletion rules were easy to apply to these simple texts and, therefore, performance was nearly perfect before, during and after training for both the Normal and Remedial Writers. Performance on the superordination rule was very good after only minimal instruction and all students regardless of ability learned to use it well.

Performance on the selection rule was never as good as performance on the superordination rule. Nevertheless, after students were taught to select topic sentences they continued to do so at a reasonable level of proficiency at least two weeks after training had ceased. Average writers were more adept at using the rule, but all students seemed to try. Finally, the invention rule was the most difficult for all subjects. Training was helpful and did result in lasting improvement but it took the very explicit training (Rules Plus Monitoring) before the performance of the remedial students reached that set by four-year college students.

The Remedial Readers, students with the most serious problems, also showed improvement following training. They were readily prompted into using the deletion rules as effectively as average students do without training. The superordination rule was also learned easily and students maintained good use of the rule on the posttests. However, improvement in the use of the selection rule was slight and poor readers experienced major problems before, during and after training with the invention rule. Instruction in the hardest rules resulted in little improvement.

Apart from providing evidence that the basic skills of summarization can be taught, Day's study also demonstrated a form of aptitude by treatment interaction (Snow, 1980) in that the more experienced students tended to need less explicit

FIGURE 3 Junior College Students' Improvement in Selecting Topic Sentences After Training (from Day, 1980)



instruction in order to improve their performance. For example, students diagnosed as poor writers needed the very explicit training involving instruction in monitoring rule use before they reached a level of performance similar to that set by untrained four-year college students. It was often the case that average and poor writers and poor readers did not differ on the pretests but, as can be seen in Figure 3, average writers benefited much more from training. Theoretically, we claim that the students of varying ability level had different "zones of proximal development" or "regions of sensitivity to instruction" (Brown, Campione, & Day, 1981; Brown & Ferrara, in press; Brown & French, 1979; Vygotsky, 1978). That is "readiness" to benefit from instruction differentiated the groups. Note that for these students the starting level of competence was approximately equal; that is, initial performance level would not differentiate the groups. However, the zone of potential development, or how much can be achieved with aid, is higher for the average than for the remedial students.

A second major finding in the Day study concerns the "explicitness" of instruction. The best performance was always obtained in the rules plus monitoring condition; the next best was the rules plus self-management group; the third best was the rules alone training; and insignificant gains were made by the self-management alone group. This ordering was found on training, on immediate and on delayed posttests. And it was obtained across rules. The consistency of this result is impressive.

In summary, training attempts were successful at improving the summarizing abilities of junior college students. The difficulty of the rules being taught and the ability level of the students interacted to determine how explicit training must be to effect improvement. In the absence of specific means to accomplish the goals, instruction of only very general self-management routines was not beneficial. Straightforward training in the specific strategies needed for problem solution led to better performance on the simpler rules. However, with the more difficult concepts, particularly with students with diagnosed reading and writing problems, explicit training in strategies for accomplishing the task coupled with routines to oversee and monitor their success, was clearly the best approach.

As a final note, we would like to reiterate that the rules of summarization that we have been studying are *general* comprehension strategies, not just a set of *specific* rules for summarizing per se (Kintsch & van Dijk, 1978). These rules are the essential operations that are basic to comprehension and it is varients of these rules that students need to make explicit as procedures for studying. Making the macrorules explicit, as we have done in our training procedures, is a generally useful study skill invaluable for improving one's ability to monitor one's own comprehension. For example, checking oneself while reading to see if one has pinpointed important material and discarded trivial information, to see if one can either identify an author's topic sentence or invent one if necessary, are all checks that comprehension is progressing smoothly. If, for example, one cannot produce an adequate synopsis of a paragraph (invention rule), this is a clear sign that comprehension is *not* proceeding smoothly and remedial action is called for (Baker & Brown, in press). Given the general applicability of the basic macrorules for a wide variety of study monitoring tasks and our initial success in training rudimentary use of the rules, we are optimistic that a general cognitive skill training program could be developed using the macrorules as cornerstones. Macrorules for comprehension (1) are important, (2) have a wide range of applicability, (3) are developmentally sensitive, (4) are educationally relevant and

(5) can be taught; all, we have argued, essential features of target skills for intensive cognitive skills training programs (Brown & Campione, 1981; Brown, Campione, & Day, 1981).

V. SELF-QUESTIONING ROUTINES AND RECIPROCAL QUESTIONING

The final study I will report has just been completed and I will give only a brief synopsis of the results. The topic is a familiar one, self-questioning while reading.

One way to facilitate learning from text during reading is to engage in self-interrogation. André and Anderson (1978-79) recently developed and tested a self-questioning study technique in which high school students were taught to locate sections of text containing important points and generate questions about them. They found that generating such questions facilitated learning better than simply reading and rereading text or making up questions without regard to important points. In addition, the training was more effective for students of lower ability, suggesting that the better students had developed effective self-questioning techniques of their own. André and Anderson suggest that self-questioning may be more effective than such passive strategies as rereading because it incorporates many metacognitive components. That is, it encourages the reader to (a) set purposes for study, (b) identify and underline important segments of the material, (c) generate questions which require comprehension of the text to be correctly answered, and (d) think of possible answers to the questions. The questioning strategy leads the student to an active monitoring of the learning activity and to the engagement of strategic action.

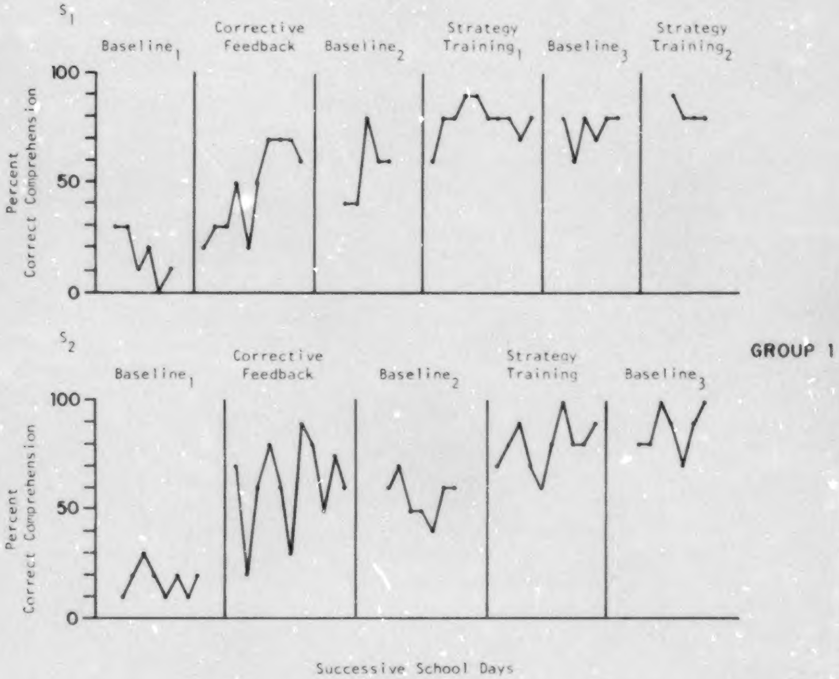
Student-generated questions are valuable not only as an aid to studying but also as an aid to comprehension itself. Singer's (1978) conception of "active comprehension" involves reacting to a text with questions and seeking answers with subsequent reading. He found that student-generated questions are more effective in promoting comprehension than teacher-generated questions even for children in elementary school. The ability to ask relevant questions of oneself during reading is, of course, crucial to comprehension monitoring and studying. Thus, training in effective question-asking may be an important first step in the development of monitoring skills. Collins, Brown, and Larkin (1980) suggest that many failures of comprehension are in fact due to a failure to ask the right questions; and a study by Nash and Torrance (1974) has shown that participation in a creative reading program designed to sensitize readers to gaps in their knowledge, such as inconsistencies and ambiguities, led to a significant improvement in the kinds of questions that first graders asked about their reading material.

A student in my laboratory, Annemarie Palincsar (Note 3), undertook to train four seventh grade minority children who were reported by their teachers to be adequate decoders (fifth-seventh grade) but poor comprehenders (second grade). The design of the study was influenced by the technology of applied behavior analysis. Palincsar had two interventions that she referred to as (1) corrective feedback and (2) strategy training. *Corrective feedback* consisted of asking children questions about a story they had read, and providing explicit information concerning the correct answer, the location of the answer in the text, and whether or not the material was implicitly (need to be inferred) or explicitly stated (Raphael, 1980). *Strategy training* (reciprocal questioning) was based on the *Re Quest* procedure developed by Manzo (1968). The student and the instructor take turns asking each other questions concerning the text. Palincsar concentrated on two of the types of comprehension fostering questions suggested by Collins and Smith (1981): hypothesis testing and prediction making. The instructor in this procedure both encourages the students to ask questions and, in turn, models the desired question type repeatedly.

Palincsar personally tested the children every day for approximately 40 days. During baseline (6-8 days), the students read grade appropriate passages and then answered the comprehension questions on the text. Each day throughout the study, they continued to read a text privately and answer 10 questions; and it is the performance of two of the students on these texts that is shown in Figure 4. In the training session, however, performance on the privately read text was preceded by either the corrective feedback or the strategy training session.

During baseline, the students achieved a level of approximately 15% correct. Following baseline, the students in corrective feedback improved to a level of approximately 50% correct, although their performance was very variable. They maintained this level of performance in the second baseline (sessions with no training). When the strategy training (reciprocal questioning) was introduced, performance soared to an impressive 80-90% correct level. Remember that these scores, shown in detail in Figure 4, were obtained on the privately read passages, i.e., different texts the students read after their interaction with the

FIGURE 4 Improvement on Comprehension Measures During Baseline and Training Sessions (from Palincsar, Note 3)



instructor. What was learned during the instructional sequence was used independently by the learners, both in the laboratory setting and in subsequent classroom behavior (Palincsar, Note 3). Modelling in how to ask appropriate questions greatly increased the students' independent comprehension scores, to the delight of their teachers.

We are greatly encouraged by the success of the reciprocal training procedure and are currently developing a version to use with younger, or less mature, coders (Brown, Echols, & Palincsar, work in progress). We begin by introducing the question game as a "fun activity" outside the context of reading. The student and the tutor watch a video-film with predetermined interest to young children of the popular suspense-surprise mode (Lichtenstein & Brewer, 1980). The film is interrupted at preset (suspenseful) intervals and the students and the teacher take turns asking each other questions. The teacher models three types of questions (inferential reasoning, analogy seeking, and prediction and hypotheses generation). Students' questions are systematically evaluated to serve as an indication of their starting competence; changes over sessions indicate their susceptibility of the model's influence. If students persist in asking inappropriate questions (such as literal facts only), the tutor, on subsequent sessions, instigates explicit training by pointing out what kinds of questions she wants and how they help comprehension. Following training, trained students will be paired with naive peers and asked to play the reciprocal questioning game with them. Such a procedure should demonstrate what the trainee has learned.

After practice with the films, students are asked to use the technique during reading sessions (if and only if there was an adequate level of questioning during the film sessions). Student and tutor again question each other at predetermined points of the text. The performance of film-trained students will be compared to a matched group of students who did not receive this treatment.

Immediately following each interactive session, the students receive a comparable text and are asked to study it on their own—underlining and writing down questions for themselves if they so wish. Following private study, students are asked to answer comprehension questions on both the interacted on and private texts. These measures of retention and comprehension will reveal a great deal about the effects of training.

Again, following training on the reading task, we plan to have trained students play the reciprocal questioning game with untrained peers. We have prior success using this method as an index of what children have gained from training (Brown, 1980; Brown & Ferrara, in press). A nice illustration of how informative such a procedure can be was provided by Au (in preparation). She asked third grade, native Hawaiian children to give a reading lesson to naive first graders, exactly as their teacher did. The children were recidivists from a program that placed heavy emphasis on comprehension training. Good readers asked their small tutees a variety of comprehension type questions ("Why do you think the mouse helped the lion?"). Intermediate readers asked far fewer questions, mainly of the literal or text-explicit kind ("What did the lion have in his foot?"). Poor readers rarely questioned their charges at all, concentrating on rereading, decoding and pronunciation. Via such techniques it is possible to evaluate the tutor's own understanding of the reading process.

VI. CONCLUSION

This series of studies has been directed at uncovering the development of selective attention strategies for understanding and remembering prose. During the junior high and high school years, students develop and increasingly fine-tune a battery of serviceable skills for learning from texts. These include: underlining (Brown & Smiley, 1978), taking notes on main ideas (Brown & Smiley, 1978), developing macrorules for comprehension, retention, and synopsis writing (Brown & Day, Note 2; Day, 1980), outlining (Armbruster, 1979), self-questioning (André and Anderson, 1978; Palincsar, Note 3), concentrating on previously missed or difficult segments of text (Brown & Campione, 1977, 1979; Brown, Campione, & Barclay, 1979; Brown, Smiley, & Lawton, 1978), and the general propensity of treating studying as a purposive act of attention-directing and self-questioning. All these skills tend to be relatively late in developing because they require a fine degree of sensitivity to the demands of learning from texts. Learners must develop rudimentary knowledge of (1) available strategies for directing attention, (2) their own characteristics as learners, including capacity limitations and background knowledge, (3) text characteristics, including important elements, structural features, etc. and (4) the nature of the criterial task, or the test to which the learning must be put (Brown, in press a, b; Brown, Bransford, & Ferrara, in press; Brown, Campione, & Day, 1981). The complex coordination of all these factors enables the student to plan, monitor and evaluate his interaction with texts in an economical and efficient manner.

ACKNOWLEDGMENTS

The preparation of this manuscript and the research reported herein was supported by the following grants: HD 06864, HD 05951 and a Research Career Development Award HD 00111, from the National Institute of Child Health and Human Development, and under Contract No. HEW-NIE-C-400-76-0116. I would like to express my appreciation to the many students and colleagues who have collaborated in various stages of the research, notably Linda Baker, Craig Barclay, Joe Campione, Jeanne Day, Roberta Jones, Sallie Lawton, Annemarie Palincsar and Sandra Smiley.

REFERENCES

Anderson, T. H. Study strategies and aids. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension.* Hillsdale, N. J.: Erlbaum, 1980.

Anderson, T. H., & Armbruster, B. B. Studying. In P. D. Pearson (Ed.), *Handbook of reading research.* New York: Longman, 1980.

André, M. D. A., & Anderson, T. H. The development and evaluation of a self-questioning study technique. *Reading Research Quarterly,* 1978-1979, *14,* 605-623.

Armbruster, B. B. *An investigation of the effectiveness of "mapping" text as a studying strategy for middle school students.* Unpublished doctoral dissertation, University of Illinois, 1979.

Baker, L., & Brown, A. L. Metacognitive skills of reading. In P. D. Pearson (Ed.), *Handbook of reading research.* New York: Longman, in press (a).

Baker, L., & Brown, A. L. Cognitive monitoring in reading and studying. In J. Flood (Ed.), *Understanding reading comprehension.* Newark, Del.: International Reading Association, in press (b).

Beauchamp, W. A preliminary experimental study of the technique in the mastery of subject matter in elementary physical science. *Supplementary Educational Monographs, 24*, University of Chicago, 1923.

Bereiter, C., & Scardamalia, M. From conversation to composition: The role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2). Hillsdale, N. J.: Erlbaum, in press.

Bessey, M. A., & Coffin, C. P. *Reading through precis*. New York: Appleton-Century, 1934.

Binet, A., & Henri, V. La memoire des phrases (Memoire des idées). *Année Psychologique*, 1894, *1*, 24-59.

Brown, A. L. The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10). New York: Academic Press, 1975.

Brown, A. L. Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1980.

Brown, A. L. Learning to learn how to read. To appear in J. Langer & T. Smith-Burke (Eds.), *Reader meets author, bridging the gap: A psycholinguistic and social linguistic perspective*. International Reading Association. Newark, Del.: Dell Publishing, in press (a).

Brown, A. L. Learning and development: The problems of compatibility, access and induction. *Human Development*, in press (b).

Brown, A. L. Knowing when, where, and how to remember; A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology*. Hillsdale, N. J.: Erlbaum, 1978.

Brown, A. L., Bransford, J. D., & Ferrara, R. A. Learning, understanding and remembering. To appear in J. H. Flavell & E. M. Markman (Eds.), *Handbook of development psychology (Vol. 1) Cognition*, in preparation.

Brown, A. L., & Campione, J. C. Training strategic study time apportionment in educable retarded children. *Intelligence*, 1977, *1*, 94-107.

Brown, A. L., & Campione, J. C. The effects of knowledge and experience on the formation of retrieval plans for studying from texts. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press, 1979.

Brown, A. L., & Campione, J. C. Inducing flexible thinking: A problem of access. In M. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning*. New York: Plenum Press, 1981.

Brown, A. L., Campione, J. C., & Barclay, C. R. Training self-checking routines for estimating test readiness: Generalization from list learning to prose recall. *Child Development*, 1979, *50*, 501-512.

Brown, A. L., Campione, J. C., & Day, J. D. Learning to learn: On training students to learn from texts. *Educational Researcher*, 1981, *10* (2), 14-21.

Brown, A. L., & Ferrara, R. A. Diagnosing zones of proximal development: An alternative to standardized testing? To appear in J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives*. Boston: Cambridge University Press, in press.

Brown, A. L., & French, L. A. The zone of potential development: Implications for intelligence testing in the year 2000. *Intelligence*, 1979, *3*, 255-277.

Brown, A. L., & Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, 1977, *48*, 1-8.

Brown, A. L., & Smiley, S. S. The development of strategies for studying texts. *Child Development*, 1978, *49*, 1076-1088.

Brown, A. L., & Smiley, S. S., Day, J. D., Townsend, M., & Lawton, S. C. Intrusion of a thematic idea in children's recall of prose. *Child Development*, 1977, *48*, 1454-1466.

Brown, A. L., Smiley, S. S., & Lawton, S. C. The effects of experience on the selection of suitable retrieval cues for studying texts. *Child Development*, 1978, *49*, 829-835.

Collins, A., Brown, J. S., & Larkin, K. M. Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1980.

Collins, A., & Smith, E. E. Teaching the process of reading comprehension. *Intelligence*, 1981.

Day, J. D. *Training summarization skills: A comparison of teaching methods*. Unpublished doctoral dissertation, University of Illinois, 1980.

Dewey, J. *How we think*. Boston: Heath, 1910 (2nd ed., Heath, 1933). In D. K. Dettermon and R. J. Sternberg (Eds.), How and how much can intelligence be increased? Norwood, N. J.: Ablex, in press.

Dijk, T. A. van, & Kintsch, W. Cognitive psychology and discourse: Recalling and summarizing stories. In W. U. Dressler (Ed.), *Trends in text linguistics*. New York: de Gruyter, 1978.

Dynes, J. J. Comparison of two methods of studying history. *Journal of Experimental Education.* 1932, *1*, 42-45.

Germane, C. E. Outlining and summarizing compared with rereading as methods of studying. In G. Whipple (Ed.), *The Twentieth Yearbook of the National Society for the Study of Education.* Part II 1921a.

Germane, C. E. Value of the written paragraph summary. *Journal of Educational Research*, 1921b, 3, 116-123.

Hasher, L., & Zacks, R. T. Automatic and effortless processes in memory. *Journal of Experimental Psychology General:*, 1979, *108*, 356-388.

Huey, E. B. *The psychology and pedagogy of reading*. Cambridge, Mass.: The M. I. T. Press, 1968.

Johnson, N. *A structural analysis of the development of story recall and summarization*. Unpublished doctoral thesis. University of California, San Diego, 1978.

Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. *Psychological Review*, 1978, *85*, 363-394.

Lichtenstein, E. H., & Brewer, W. F. Memory for goal-directed events. *Cognitive Psychology*, 1980, *12*, 412-445.

Manzo, A. V. *Improving reading comprehension through reciprocal questioning*. Unpublished doctoral dissertation, Syracuse University, 1968.

Masur, E. F., McIntyre, C. W., & Flavell, J. H. Development changes in apportionment of study time among items in a multitrial free recall task. *Journal of Experimental Child Psychology*, 1973, *15*, 237-246.

McClusky, F. D., & Dolch, E. W. A study outline test. *School Review*, 1924, *34*, 757-772.

Nash, W. A., & Torrance, E. P. Creative reading and the questioning abilities of young children. *Journal of Creative Behavior*, 1974, *8*, 15-19.

Perry, W. Students' use and misuse of reading skills: A report to the faculty. *Harvard Educational Review*, 1959, *29*, 193-200.

Piaget, J. *Judgment and reasoning in the child*. New York: Harcourt, 1928.

Preston, R., & Tufts, E. The reading habits of superior college students. *Journal of Experimental Education*, 1948, *16*, 196-202.

Raphael, T. E. *The effects of metacognitive strategy awareness training on students' question answering behavior*. Unpublished doctoral thesis. University of Illinois, 1980.

Scardamalia, M., & Bereiter, C. *Fostering the development of self-regulation in children's knowledge processing*. Paper presented at NIE-LRDC Conference on Thinking and Learning Skills, Pittsburgh, October, 1980.

Singer, H. Active comprehension from answering to asking questions. *Reading Teacher*, 1978, *31*, 901-908.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Journal of Educational Psychology* , 1977, *69*, 381-387.

Snow, R. E. Aptitude Processes. In R. E. Snow, P. A. Federico, & W. E. Montague (Eds.), *Aptitude, learning, and instruction* (Vol. 1). Hillsdale, N. J.: Erlbaum, 1980.

Stordahl, K. E., & Christensen, C. M. The effect of study techniques on comprehension and retention. *Journal of Educational Research*, 1956, *49*, 561-570.

Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, 1917, *8*, 323-332.

Vygotsky, L. S. *Mind in society: The development of higher psychological processes.* (Eds.) M. Cole, V. John-Steiner, S. Scribner, & E. Souberman. Cambridge, Mass.: Harvard University Press, 1978.

Weinstein, C. Elaboration skills as a learning strategy. In H. F. O'Neil (Ed.), *Learning strategies*. New York: Academic Press, 1978.

REFERENCE NOTES

1. Collins, A., Brown, A. L., Morgan, J. L., & Brewer, W. F. The analysis of reading tasks and texts. Center for the Study of Reading Technical Report, University of Illinois, 1977.
2. Brown, A. L., & Day, J. D. Strategies and knowledge for summarizing texts: The development of expertise. Unpublished manuscript, University of Illinois, 1980.
3. Palincsar, A. Corrective feedback and strategy training to improve the comprehension of poor readers. Unpublished manuscript, University of Illinois, 1981.

TAFFY E. RAPHAEL
University of Utah

ROBERT J. TIERNEY
University of Illinois
Center for the Study of Reading

THE INFLUENCE OF TOPIC FAMILIARITY AND THE AUTHOR-READER RELATIONSHIP ON DETECTION OF INCONSISTENT INFORMATION

Research related to how written discourse is understood has been concerned with reader and text variables as they affect memory for and understanding of text. Within the focus of reader variables lie those studies of "metacomprehension," of one's *awareness of* and *control over* the process behind discourse comprehension.

The notion of "control" in metacomprehension involves the reader's knowledge of his or her ability to both regulate and monitor on-going reading behavior, whether it be decoding or comprehension (Flavell, 1976). A commonly used method for examining this knowledge has been the error detection paradigm. This has involved such tasks as having young students respond to incomplete instructions (Markman, 1977) and having college students or young children respond to passages containing inconsistent information (Baker, 1979; Winograd & Johnston, 1980). With few exceptions (e.g., Garner, 1980), across tasks and age groups, results have indicated that the readers and/or listeners are consistently unaware of errors in passages.

However, it may be premature to conclude as a result of these studies that readers have not understood the text or processed the error. Rather, many plausible reasons exist to explain the consistent lack of error detection. Baker (1979) suggests several such as the reader consciously decides the errors are trivial, hence not worth mentioning in free recall; or the reader selectively omits and/or transforms the information to make sense of the text. Winograd and Johnston offer several additional reasons. For example, they suggest readers may have an overriding faith in the Cooperative Principle described by Grice (1975), a principle which suggests that both speakers and writers usually intend to make their message truthful, relevant, and unambiguous. This Cooperative Principle may cause the reader to draw inferences to make the inconsistent information meaningful. The evidence that readers do not recognize inconsistent information in text exists. Hence, rather than question *that* it occurs, the more interesting questions now appear to be *when* and *why* inconsistent information is either accepted or ignored.

Of the factors which can influence the reader's awareness of and reaction to a passage, topic familiarity, discourse style and text structure were selected for manipulation in this study: (1) topic familiarity, documented as having a major influence on discourse comprehension (Bransford & Johnson, 1972; Anderson, Reynolds, Schallert & Goet, 1978), (2) discourse style, representative of the relationship between the author and the reader (Brewer, 1981), and (3) text structure, which affects understanding in terms of the "match" between the expectations and hypotheses generated by the reader and the information presented by the text (Thorndyke, 1977).

Studies utilizing the error detection paradigm have not systematically investigated the effects of these variables, and have generally relied heavily on free recall measures and question probes in assessing the effect of the inconsistencies on comprehension (e.g., Markman, 1977). The current study extends this work, first by attempting to isolate the effects of topic familiarity, discourse style, and structural variations; and second by adding an "interview" assessment geared toward identifying specific reading behaviors which occur at the point of the inconsistent information. Therefore, the purposes of this study were to (1) examine the effect of inconsistent information on fifth grade students' prose comprehension as measured by the inclusion of the inconsistency in free recall, (2) examine those reading behaviors that occur during the time the reader is faced with the inconsistent information, and (3) examine these effects and behaviors across texts which vary on the dimensions of familiarity, discourse style, and placement of errors in the hierarchical organization determined for the text.

METHOD

Subjects

Forty-three fifth grade students from a central Illinois public school participated in the study. Students represented a range of ability levels and varied on socio-economic status, sex, and ethnic background. Students were randomly assigned to one of the eight conditions created for purposes of counterbalancing.

Materials

Sixteen passages based upon four different themes were used in the study. Versions of two of the themes varied on the dimension of topic familiarity. Familiarity was defined in terms of continuum from abstract to concrete (Thorndyke, 1977), with corporate structures deemed relatively unfamiliar to fifth grade students, and children at play deemed a familiar, concrete situation (Raphael, Myers, Freebody, Tirre & Fritz, 1980). Data from previous studies using these passages supported our familiarity distinction. Versions of the remaining two themes varied on the dimension of discourse style. Discourse style can be considered to be that dimension concerned with the author's selected method of presentation. In this study, the dimension related to the use of dialogue to transmit information was selected for manipulation. Thus, there were four passages for each of the two manipulations. Excerpts from two passages, one from each manipulation, can be seen in Table 1. Inconsistent information was then created for each passage at three levels in the structure determined for the text: superordinate or general, subordinate or detail, and author-related or personna (see Table 2). A total of six inconsistent statements were created, two each at the superordinate, subordinate, and author-related levels. In creating the final versions, each text had two versions containing six target points—three of inconsistent information (one subordinate, superordinate, and author-related) and three consistent target points. In the comparable passage version, the opposite in terms of consistent/inconsistent target points existed.

Passages were approximately 375 words in length and ranged in readability levels from 3.5 to 5.0. Pilot data indicated that the error free versions of all passages were comprehendable by average fifth grade students.

TABLE 1. Passage Manipulation

Familiarity	
Familiar	**Less Familiar**
Mary has a problem. This is the story of how *she solved* it (could not solve it). There were two things that Mary had always wanted. One was a place to be alone and the other was a place that she could use for bird watching. Mary decided that there was one way she could get both of these things. Her family was renting a home in the big city, and in the back yard, away from the house, they had a large tree. She made up her mind to build a treehouse in that tree. That way she could do the things she wanted to do by herself. . .	The Poly Plastic Bag Company had a problem. This is the story of how they *solved it* (could not solve it). There were two things that the Poly Plastic Bag Company had always wanted. One was a factory of its own and the other was offices that were out of the city. They decided that they could get both of these two things. They were currently renting a factory in a big city. But near a quiet river out of the city, they owned a large block of land. They made up their minds to build a factory on that land. . .
Discourse Style	
Non-Dialogue	**Dialogue**
All over the world children *like* (hate) to play different games. In some countries, children enjoy playing a game called "Fly." It gets its name because to play the game you need to be able to leap through the air. Some people say that it is like flying through the air. The game is very easy to learn and play. The only equipment you need to have is six sticks that are similar in size. The sticks need to be as long as a person's foot and about as big around as a person's thumb. After the sticks are found, they are placed on the ground. . .	Lisa and Mike were bored. It was Saturday, and they did not know what to do until Lisa had an idea. "I know a game we can play that they play in some countries. You know children all over the world *like* (hate) to play different games," Lisa said. Mike was interested and asked, "What is the game?" "It is called Fly because to play the game you need to be able to leap through the air. Some people say you have to fly through the air," Lisa said. "You only need six sticks to play it." It sounded like fun so they decided to try the game. Lisa and Mike gathered six sticks, each one about as long as their feet and as big around as their thumbs. They placed the six sticks on the ground. . .

Procedure

Students met individually with one of four experimenters for 45-70 minute sessions. The procedure used with all students involved the following steps for four passages: (1) reading the passage, (2) providing a free recall, (3) responding to probe questions dealing with nontarget information, and (4) oral reading of the passage, which included an interview at the six target points in the passage (three

TABLE 2. Embedded Inconsistencies

	CATEGORY OF INCONSISTENCY		
PASSAGE	**Author-related**	**Subordinate**	**Superordinate**
Familiar	Mary had a problem. This is the story of how she *solved* (could not solve) it.	Immediately her parents said the building would have to *stop* (be finished more quickly).	Things turned out even *better* (worse) for Mary.
Unfamiliar	The Poly Plastic Bag Company had a problem. This is the story of how they *solved* (could not solve) it.	Immediately, the Pollution Control Board said the building would have to *stop* (be finished more quickly).	Things turned out even *better* (worse) for the company.
Dialogue	". . . You know children all over the world *like* (hate) to play different games," Lisa said.	". . . You have to go from one end to the other by *stepping only one foot* (singing only one song) in each space, never touching a stick . . ."	". . . If you touch (a stick) you're *out* (winner) of the game."
Non-Dialogue	All over the world children like (hate) to play different games.	He then steps across the first stick, *placing a foot* (singing a song) in the first space.	If a player does touch a stick, he is *out* (winner) of the game.

at points of information, three at non-altered points). This was repeated for four passages, and finally followed by a general debriefing. In other words, students read, recalled, and answered probe questions, then orally read the passage in sections. Each section included error 1 or 2, 3 or 4, 5 or 6. Interview points occurred at the 6 points where an error *could* exist: three points presented inconsistent data, three consistent information. At these interview points students were asked such questions as "What does ________ mean?" "Does it make sense?" "What information did you use to help figure this out?" "Would you rewrite it so that it would be easier to understand?" This interview was recorded and later transcribed.

Scoring

Pertinent to the purposes of the study were the free recall and the interview data. Initial scoring of the free recall data involved general identification of inconsistencies as (1) omitted from recall, (2) included in recall and resolved, and (3) included in recall and unresolved.

Scoring the interview data involved categorizing responses into one of thirteen categories. These discrete categories were designed to represent the range of reading behaviors a student could perform in coping with the inconsistency. For example, the error could be unrecognized as an error for reasons such as (1) misreading the text, (2) reading the text correctly but drawing an inference that would make the inconsistency meaningful, or (3) reading the text correctly and truly missing the inconsistency.

RESULTS AND DISCUSSION

The following questions guided the examination of these data: (1) Did the readers' free recalls include inconsistent information? (2) Did th[illegible]terview provide additional information regarding comprehension monitorin[illegible] Were the methods used by the reader to resolve inconsistencies differen[illegible]oss the four text conditions (familiar vs. less familiar, dialogue vs. non-dialogue) and error types (superordinate, subordinate, and author-related)?

Frequency data for the free recalls and the interviews provided answers to these questions. Whereas the interview data indicated that most readers had addressed the inconsistencies and were monitoring comprehension, the inconsistencies were rarely included in the free recall data. The interviews indicated that over 90 percent of the readers either had developed a well-reasoned account which "made sense" of each inconsistency or had recognized the inconsistencies as errors. In contrast, fewer than 10 percent of the readers' free recalls included reference to any item which involved inconsistent information.

Chi square analysis was the basis for examining whether the methods used to resolve inconsistent information varied across error types as well as within and across text conditions. Within the familiar, the dialogue and the non-dialogue text conditions, significant chi square statistics were obtained across error type (X^2 = 38.56, 41.52, 71.78; d.f. = 24; $p < .01$). That is, within three of the four text conditions, readers varied in the method they used to resolve different types of inconsistent information. In the fourth text condition (less familiar) there were no significant differences.

Differences across text conditions were less substantial. Across dialogue and non-dialogue and familiar and unfamiliar text conditions, significant differences existed across selected methods of resolution. For example, across dialogue conditions, readers in the non-dialogue condition rarely assumed an indirect speech act when faced with an author-related inconsistency. However, in the dialogue condition, this method of resolution was often invoked. Among the differences that existed across familiarity conditions was the reader's indication of the knowledge of events. In the familiar condition, when faced with a superordinate or an author-related inconsistency, readers suggested the need for an author to add specific information. In the unfamiliar condition no students used this method of resolution except in response to inconsistencies on a subordinate level. Table 2 details these patterns of differences across text conditions for each of the various methods of resolution.

The results of this study are consistent with those findings of Markman (1977), Baker (1979), and Winograd and Johnston (1980): readers did not include inconsistent information in their free recall, as indicated by the free recall data. However, these data are in sharp contrast to the interview data. The interview data suggests readers do monitor their comprehension but that this monitoring would not be exposed unless probed. Furthermore, the interview data provides information concerning *why* the errors were not apparently recognized. The data substantiates suggestions made by both Baker (1977) and Winograd and

Johnston (1980): (1) readers may feel the information is trivial, or obviously incorrect, and is thus not worth including in free recall, and (2) the Grician (Cooperative Principle) is operating in many cases such that the reader creates a script-based scenario or does some text integration to make sense of the inconistency. In addition, the readers' expectations are often strong enough to elicit a semantically appropriate miscue, thus changing inconsistent information to the more predictable form.

Finally, the data suggests that the error detection paradigm can reveal much about the readers' behavior when forced into situations requiring them to monitor their comprehension, and that the free recall data should *not* be interpreted to mean readers do not or are incapable of monitoring their comprehension.

IMPLICATIONS OF THE STUDY

The implications which can be drawn from this study fall into three categories: methodological, instructional, and psychological. The first addresses the use and appropriateness of the error detection paradigm. It appears that much can be learned of the readers' cognitive behavior during reading by forcing the reader into a "debugging" state through inserting errors. However, free recall is not the optimal means for assessment. The interview procedure allowed the identification of a variety of strategies for resolving inconsistent information which suggests that readers did not enter a purely editorial mode when asked about the error.

The second implication represents instructional concerns. Students were able to comment upon (thus we infer, were aware of) reading strategies when asked only one or two questions. This suggests that it may be possible to induce comprehension monitoring behaviors with a minimal amount of instruction.

Finally, the psychological implication addresses issues for further research. Rather than a blanket statement as to ability of readers to detect errors in text, the more interesting issue is to realize the complexity of the cognitive operations driven by the pragmatics of discourse—an issue which this study has just begun to explore.

REFERENCES

Anderson, R. C., Reynolds, R. E., Schallert, D. L., and Goetz, E. T. Frameworks for comprehending discourse. *American Educational Research Journal*, 1977, *14*, 367-382.

Baker, L. *Comprehension monitoring: Identifying and coping with text confusions*. Technical Report 145, University of Illinois Center for the Study of Reading, September, 1979.

Bransford, J. D. & Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 717-726.

Brewer, W. F. Literary theory, rhetoric, stylistics: Implications for psychology. In R. Spiro, B. C. Bruce and W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1981.

Flavell, J. H. Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, N. J.: Lawrence Erlbaum and Associates, 1976.

Garner, R. Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. *Journal of Reading Behavior*, 1981, *12*, (1), 55-63.

Grice, H. P. Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics, 3, Speech acts*. New York: Seminar Press, 1975, 41-58.

Markman, E. M. Realizing that you don't understand: A preliminary investigation. *Child Development*, 1977, *48*, 986-992.

Raphael, T. E., Myers, A. C., Freebody, P., Tirre, W. C., & Fritz, M. *Contrasting the effects of some text variables on comprehension and ratings of comprehensiblity.* (Technical Report 190.) Urbana, Ill.: University of Illinois, Center for the Study of Reading, December, 1980.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, 9, 77-110.

Winograd, P. & Johnston, P. *Comprehension Monitoring and the Error Detection Paradigm.* Technical Report 153, University of Illinois, Center for the Study of Reading, January, 1980.

WAYNE L. SHEBILSKE
University of Virginia

DENNIS F. FISHER
Human Engineering Laboratory,
Aberdeen Proving Ground, MD.

EYE MOVEMENTS REVEAL COMPONENTS OF FLEXIBLE READING STRATEGIES

When asked to characterize superior readers, educators single out flexible reading, "the ability to adjust one's rate and approach to reading with the purpose of reading, with the difficulty of material, and with one's background or knowledge of the particular subject matter" (Weintraub, 1977, p. 169; see also Fisher & Montanary, 1977; Harris & Sipay, 1975; and Tinker, 1947). Despite its importance, very little is known about adaptable reading strategies and as a result, educators have not developed effective lesson plans for teaching it. One reason is that flexibility has been treated as a modulating single strategy when it is in fact a compound of fundamental component strategies representing successive stages in the acquisition process. In the present paper we use eye movement records to analyze and parse out components of flexible reading.

Some researchers have examined flexibility as a between passage strategy using total reading time over an entire passage as their main response measure where typography and content were varied (cf. Fisher, 1976; Rankin, 1974). We, on the other hand, follow the lead of Shebilske and Reid (1979) who used eye movement records to analyze components of flexible reading within passages. Shebilske and Reid based their analysis on a reading model according to which eye movements are controlled by recognition and comprehension. *Recognition* processes map light patterns onto surface structural units corresponding to letters, syllables, sentences, and paragraphs. (Several recent reaffirmations of the influence of recognition processes on eye movements have been made (cf. Just and Carpenter, 1980; Rayner, 1978). *Comprehension* processes map the output from recognition processes onto conceptual representations and build conceptual units into structured internalized representations.

Shebilske and Reid provided support that the comprehension processes influence eye movements by showing that students reading a 2,000 word narrative slow down when reading sentences that are indirectly related to higher-order conceptual units than they do over sentences that are directly related, but when these same sentences were read out of context, no difference was found. They argued that their contextual effects represented components of intra-passage flexibility. When they compared intra- with inter-passage flexibility they found a significant positive correlation. The present study extends eye movement measures of the hypothesized flexibility and adaptability to textbook reading.

Two problems must be solved in order to measure components of flexible reading: 1) defining units of analysis, and 2) specifying relevant characteristics of those units. The acid test for any recommended solution is whether or not the units and characteristics can be used to shed light on adaptable reading. In the present study we tested units and their characteristics that were developed by a multidisciplinary team effort. (Members of this team were James Deese, John

Rotundo, Thomas Estes and Mary Wetmore in addition to the authors.) The units and characteristics were defined from the point of view that what is read is as important as what an author has written and therefore the ways ideas are segmented, and characteristics like importance and familiarity are expected to be revealed in changes in eye movement dynamics like fixation duration, frequency and distance between subsequent fixations.

METHODS

Subjects. Two technical staff employees at the Human Engineering Laboratory volunteered to participate in our experiment. Both had received undergraduate college degrees in areas other than biology.

Apparatus and Materials. We measured eye movements with the oculometer located at the Behavioral Research Directorate, Human Engineering Laboratory, at Aberdeen Proving Ground, Maryland. Further details on this free viewing system are reported by Monty (1975). Reading material was projected onto a 24″ × 40″ screen located 1.7 meters in front of readers who sat in an easy chair without head restraints. Since all recording equipment was located in a separate room, subjects were not aware that their eye movements were being monitored. Calibration was accomplished in tasks that did not give away its true purpose. Horizontal eye position was measured with an accuracy of approximately 1° (4 letter word), and vertical eye position was measured with an accuracy of approximately 1° (one line). This TV based monitoring system utilized PDP 11/20 and PDP 11/50 computers for pupil/iris acquisition and on-line analysis. A Tektronix hard copy unit provides scan-patterns of segmented fixations.

The text was a 2,866 word excerpt from a tenth grade biology textbook (*BSCS,* Green Version, 2nd edition), typed in pica and double spaced with about 15 lines per page.

Procedure. Subjects were told to read the excerpt as they would ordinarily read such a passage for a homework assignment. They proceeded at their own pace pushing a button to advance slides to new pages. They were told that they would be tested with detailed essay and multiple choice questions and that they would have an opportunity to reread the text before the test. Tests of paraphrastic recall and multiple choice followed the second reading.

RESULTS

Several preliminary analyses assessed the representativeness of our two readers. Recalls proved comparable to those given by University of Virginia undergraduates in our previous experiments and that reading rates were comparable to college student norms. One reader averaged 232 words per minute on the first reading and the other averaged 335 words per minute. The average reading rates for college students is about 280 words per minute for easy material (Taylor, Frankenpohl, & Pette, 1960).

Scan paths and digital records, like those in Figure 1, of critical parameters of each fixation like duration, location and distance from previous fixation were analyzed. Each reader produced more than twelve thousand data points for each reading. To overcome a potential overwhelming problem a hierarchical method of data analysis was devised. First analyzed were macro-level aggregates of eye movements, then the micro-parameters of individual fixations.

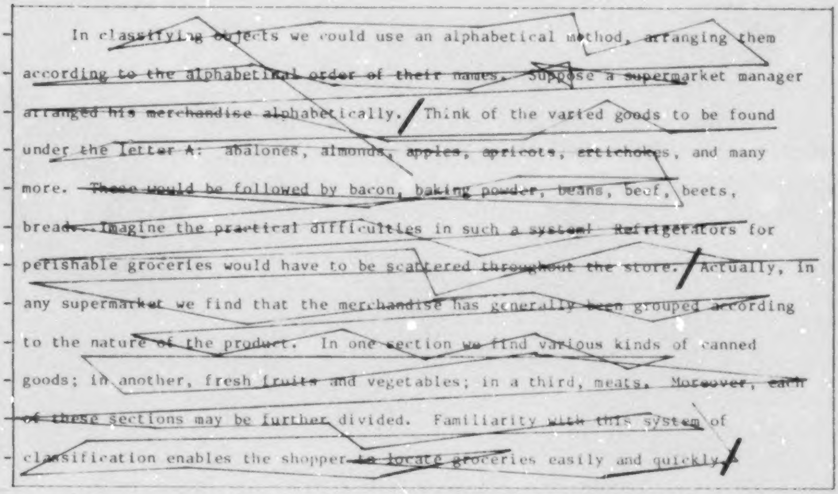
The macro-level analysis measured modulations in reading rate as a function of meaning unit characteristics. The meaning units were defined by ratings of importance made by hundreds of University of Virginia students on segments of a 2,500 word text. Three such units are shown in Figure 1. For each meaning unit three scores were averaged over both subjects: reading rates for the first and second readings, and the difference in rate between the first and second reading. These data were then correlated with the first reading rate and the difference rate with the 18 meaning unit characteristics shown in Table 1.

The results of the analysis of the first reading data showed that subjects slowed down for ideas that tended to be recalled ($r = -.12$, $p = <.15$), for important ideas ($r = -.17$, $p = <.07$), for ideas that contained new or unfamiliar information ($r = .22$, $p = <.03$), and for ideas containing a high number of propositions essential to the gist of that idea ($r = .11$, $p = <.17$).

Analyses of the changes between first and second readings showed that the difference rate was correlated with meaning unit importance ratings ($r = .44$, $p<.001$), the average importance of propositions in an idea ($r = .50$, $p<.001$), the propositions essential to the gist of an idea ($r = .30$, $p<.01$), and serial position ($r = -.28$, $p<.01$). The correlation with serial position indicates that on a second reading subjects slowed down more towards the end of the passage. One possible explanation for this is that there was more important information near the end of this passage.

The overall pattern of correlations show that readers modulate their rate in accord with the familiarity (old or new) of information on the first reading, spending more time on new information, and to whether information is viewed as important or unimportant on the second reading. The mean reading rates for important vs.

Figure 1. Eye movement scan paths obtained from a reader on a portion of text with three meaning units.



In classifying objects we could use an alphabetical method, arranging them
according to the alphabetical order of their names. Suppose a supermarket manager
arranged his merchandise alphabetically. Think of the varied goods to be found
under the letter A: abalones, almonds, apples, apricots, artichokes, and many
more. These would be followed by bacon, baking powder, beans, beef, beets,
bread. Imagine the practical difficulties in such a system. Refrigerators for
perishable groceries would have to be scattered throughout the store. Actually, in
any supermarket we find that the merchandise has generally been grouped according
to the nature of the product. In one section we find various kinds of canned
goods; in another, fresh fruits and vegetables; in a third, meats. Moreover, each
of these sections may be further divided. Familiarity with this system of
classification enables the shopper to locate groceries easily and quickly.

TABLE 1
MEANING UNIT CHARACTERISTICS

1. MU Recall = Proportion of Subjects Recalling MU.
2. MU Importance Rating.
3. Average Importance of Propositions in MU.
4. Number of Propositions in MU.
5. Number of Words in MU.
6. Number of Propositions Essential To Gist.
7. Number of Important Propositions.
8. Proportion of Propositions Essential to Gist.
9. Proportion of Important Propositions.
10. Serial Position in Text.
11. Serial Position from Beginning of Section.
12. Serial Position from End of Section.
13. Serial Position from Beginning of Paragraph.
14. Serial Position from End of Paragraph.
15. Number of Zero-Diameter Clusters Contained in MU.
16. Proportion of Subjects Placing Segment Boundary at Beginning of MU.
17. Proportion of Subjects Placing Segment Boundary at End of MU.
18. Old-New Rating.

unimportant ideas are of special interest. On the first reading, the subjects read important and unimportant ideas at rates of 280 wpm and 288 wpm respectively. On the second reading, they read important ideas at the rate of 229 wpm and unimportant ideas at the rate of 372 wpm. That is, these subjects read important ideas 51 wpm slower and unimportant ideas 84 wpm faster than they did on the first reading. Even with two subjects, the test of this interaction with a *t* test with idea units as the random variable, resulted in a positively biased but significant, *t* (72) = 3.60, $p<.001$, effect giving power to our finding.

Since importance was the most influential meaning unit characteristic in our macro-analysis of aggregated eye movements, a micro-analyses of individual eye fixations comparing fixations on important and unimportant ideas followed. Differences in fixation duration, number of regressions, and interfixation distance were analyzed. On first and second readings similar trends appeared. Readers had longer fixations and more regressions on important ideas while interfixation distances were relatively constant. More specifically: 1) mean fixation durations of 276 msec. for important ideas vs. 240 msec. for unimportant, 2) 4 regressions per important unit vs. 1 regression per unimportant unit, and 3) average interfixation distances of 4.28 degrees for important ideas vs. 4.46 degrees for unimportant ones. It was not infrequently that readers finished a meaning unit before deciding to slow down and take another look at it. This verification process was also evident when they came to the last word in a unit (most likely an important unit) and then made a large regression, sometimes back to the first word in the idea. This result not only reveals an interesting reading strategy and provides support for comprehension effects, but also adds construct validity to our meaning units.

DISCUSSION

The data reported support the notion that macro and micro variations in eye movement patterns result from flexible reading strategies under voluntary control. At times the variations correspond to and seem determined by the text, while at other times variations reflect the *subject's* view of important information. That is, subjects could have read important and unfamilar material faster if they had chosen to do so. This interpretation will be further tested in future experiments that will manipulate variables like the purpose for reading, background knowledge, text redundancy, and end of sentence/idea unit integration time.

The prospects for future research are especially bright considering that adaptability and generality of our reader-based procedures for utilizing and characterizing reading in applied content area settings. Karmiohl (1979) recently showed that reader-based procedures can detect variations in perceived text structure that occur in a wide range of contexts like reading a passage on marketing to business students or to psychology students. Recent preliminary studies using Karmiohl's materials showed that readers move their eyes differently depending upon how they perceive the text structure. The adaptability of our reader-based procedures contrast sharply with the limitations of formal text grammars being used to unitize and characterize texts in other laboratories (e.g. Just & Carpenter, 1980). Formal grammars measure only the structure that an author puts into a text; they do not measure the structures the readers see in a text (cf. Shebilske, 1980).

While a formal text grammar applies to one discourse style, others in our research group have identified and measured a variety of discourse styles in textbooks (cf. Deese, 1980). The adaptability and generality of our approach makes it particularly well suited for studying individual differences based on background. Background effects were shown in the present study when students regulated their eye movements according to whether information was old or new. Individuals also differ in their familiarity with various styles of discourse and in their ability to discover modifiers and to make inferences. The present study represents a starting point from which to move toward comprehensive analyses of individuals' performance.

Specific recommendations for teaching may be premature, but some suggestions for application are appropriate and timely. This project is based upon the critical assumption that the mechanics of reading are largely controlled by the level of understanding and not the other way around. Therefore, the path of previous attempts to improve reading by improving the mechanics of eye movements (cf. Tinker, 1958) goes counter to our purpose. For the most part, eye movement measures will remain in the laboratory and provide one of several converging operations used to infer effective, flexible reading strategies. Our goal will be to encourage the use of flexible strategies on the part of the reader and to inform the text writer and producer about the facilitory aspects of more efficient text analysis. At the classroom level we are coordinating our research efforts with teachers in an attempt to make our data a sound empirical foundation for developing effectual lesson plans for teaching flexible reading.

REFERENCES

Deese, J. *Form of the text and comprehension*. Paper presented at the 30th Annual Meeting of the National Reading Conference. San Diego, December 1980.

Fisher, D. F. Spatial factors in reading and search: The case for space. In R. A. Monty & J. W. Senders, *Eye movements and psychological processes*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1976.

Fisher, D. F., & Montanary, W. E. Spatial and contextual factors in beginning reading: Evidence for PSG-CSG complements to developing automaticity? *Memory and Cognition*, 1977, *5*, 247-251.

Harris, A. J., & Sipay, E. R. *How to increase reading ability: A guide to developmental and remedial methods*, (6th Ed.). New York: David McKay Co., Inc., 1975.

Just, M. A., & Carpenter, P. A. A theory of reading: From eye fixations to comprehension. *Psychological Review*, 1980, *87*, 329-354.

Karmiohl, C. M. *Text-based vs. reader based methods for the determination of conceptual structure and the prediction of the process and the quality of reading comprehension.* Unpublished Masters thesis, University of Virginia, 1979.

Monty, R. A. An advanced eye movement measuring and recording system. *American Psychologist*, 1975, *30*, 331-335.

Rankin, E. F. The measurement of reading flexibility: Problems and perspectives. *Reading Information Series: Where do we go?* Newark, DE: International Reading Association, 1974.

Rayner, K. Eye movements in reading and information processing. *Psychological Bulletin*, 1978, *85*, 618-660.

Shebilske, W. L. Structuring an internal representation of text: A basis for literacy. In P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds.), *Processing of visible language* (1) Vol. 2(1). New York: Plenum, 1980.

Shebilske, W. L., & Reid, L. S. Reading eye movements, macro-structure and comprehension processes. In P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds.), *Processing of visible language* (1) Vol. (1) 1. New York: Plenum, 1979.

Taylor, S. E., Frankenpohl, H., & Pette, J. L. Grade level norms for the components of the fundamental reading skill. *EDL Information and Research Bulletin No. 3*, 1960, Huntington, New York: Educational Developmental Laboratories.

Tinker, M. A. Time relations for eye movements in reading. *Journal of Educational Psychology*, 1947, *38*, 1-10.

Tinker, M. A. Recent studies of eye movements in reading. *Psychological Bulletin*, 1958, *55*, 215-231.

Weintraub, S. Research. *The Reading Teacher*, 1977, *21*, 169-173.

RICHARD VENEZKY
University of Delaware

LETTER-SOUND REGULARITY AND ORTHOGRAPHIC STRUCTURE

English orthography occupies a highly visible role both in initial reading instruction and in psychological studies on the recognition of printed words. Almost all of the historically sanctioned methods for introducing reading instruction in this country attend, sooner or later, to the complexities of translating from spelling to sound. Although programs differ in the number of letter-sound patterns they introduce and in the manner in which letter sounds are identified, they nevertheless appear to encourage the same procedure for generating sound from spelling. Concepts about the relationship of spelling to sound and about the role which letter-sound correspondences play in learning to read, although seldom clearly stated, can nevertheless be inferred from the activities and rhetoric of these reading programs.

Psychologists, on the other hand, have been somewhat more forthcoming of late about their notions on orthography in reading and, in particular, in the recognition of printed words. For the first 70 years of reading research, from the time of Cattell until the middle 1950s, orthography was never mentioned as a potential factor in word recognition. But the work of Shannon in the late 1940s on redundancy in information transmission led in the early 1950s to empirical investigations of the statistical properties of English orthography and their influence on recognition and recall of letter strings (Miller, Bruner, and Postman, 1954).

The first speculations on the role of letter-sound correspondences in word recognition came in the work of Gibson, Pick, Osser, and Hammond (1962) who found that letters of pronounceable pseudowords were more easily recalled in a full report paradigm than were letters of unpronounceable pseudowords (e.g., *dink* vs. *nkid*). The authors hypothesized that the operative unit in word perception was not the letter (or whole word), but a sequence of letters which had an invariant letter-sound relationship. However, Gibson, Shurcliff, and Yonas (1970) obtained the same results with congenitally deaf subjects, thereby demonstrating that orthographic structure alone could account for the response differences.

Letter-sound correspondences have also been utilized in word recognition models by Rubenstein, Lewis, and Rubenstein (1971) and by Smith and Spoehr (1974). However, both the paradigms employed and the assumed use of letter-sound relationships in the latter two studies differed from that in Gibson et al. (1962). Both Rubenstein et al. (1971) and Smith and Spoehr (1974) assumed that letter recognition occurs on the basis of letter features; then sounds are attached to letters or letter sequences and lexical access made on the basis of a phonological unit. Gibson et al. (1962) assumed that word recognition occurred on the basis of higher-order visual units which the reader formed through experience with letter-sound translation.

Structural variables have also continued to play a role in word recognition models, whether based on statistical properties (e.g., Mason, 1975) or linguistic ones (e.g., Massaro, Venezky, and Taylor, 1979). Spelling patterns, letter-sound

68

patterns, orthographic structure, statistical structure, and rule-governed structure represent the more prominent of the orthography-related variables postulated for the word recognition process. What these variables are, how they differ, and what role they might play in word recognition are the primary concerns of this paper. The first section is didactic, intending to unravel the mysteries of English spelling and, in particular, to distinguish the letter-sound patterning of English from statistical and rule-governed orthographic structure. The second section attends to word recognition theories, centering primarily on the notion of phonological mediation.

AN ORTHOGRAPHY PASTICHE

Letter-Sound Regularity

English spelling is frequently cursed (e.g., Lounsbury, 1909; Pitman, 1961), much less frequently blessed (e.g., Chomsky and Halle, 1968), but seldom studied seriously. The relationship of spelling to sound in English is neither the primordial chaos of the spelling reformers, nor (except in a perverse manner) the ideal underlying structure of the generative phonologists. It has at its core a set of patterns which originated from Latin orthography of the sixth and seventh centuries, but were heavily modified over a thousand year period, primarily through the influence of sound change. Intermixed with this core, like the swirls of chocolate in a marble cake, are a host of sub-patterns and singularities created through borrowings, scribal caprice, and lexicographic tampering. Nevertheless, by the beginning of the sixteenth century, the counterpoint of morphological and etymological representation with that of the phonological base was firmly established and has undergone little alteration since. The underlying structure and specific patterns in this system are described by Venezky (1967, 1970); what is described here are those features of letter-sound regularity that are relevant for the notions of *spelling patterns* (Gibson, Pick, Osser, and Hammond, 1962) and *phonological mediation* (e.g., Rubenstein, Lewis, and Rubenstein, 1971; Smith and Spoehr, 1974; Baron and Strawson, 1976).

As a first approximation, we can class the letter-sound patterns of English as *invariant, variant-predictable* and *variant-unpredictable*. In the first class are a small group of consonants (f, j, ph, q, rh, sh, v, wh, z) that have single mappings into sound, with either no or relatively few exceptions. (It is interesting to note that with the exception of *f*, these are spellings that either did not exist or were rarely used in Old English.)

In the second group, variant-predictable, are most of the remaining consonants and, depending upon one's criterion for predictability a few of the vowels. However, to claim that the various pronunciations for a spelling are predictable is to assert that there exists a set of features f_1, ..., fn such that each pronunciation can be uniquely characterized by a distinct combination of these features. For example, the set of features {initial position; non-initial, pre-stress; non-initial, post-stress} uniquely characterize the pronunciations for *x*. In initial position it is /z/ (e.g., xylophone); in medial position but before the main word stress, it is /gz/ (e.g., example); and in any position after the main word stress, it is /ks/ (e.g., *taxi, fox*).

The central issue in establishing the class of variant-predictable pronunciations is in deciding what qualifies as a feature. Position within a word, adjacent letters and sounds, and position relative to the main word stress are all detectable by native speakers. So are morphemic identification and form class. But etymology is not, in that except for a small percentage of the vocabulary, no clues are available within words or the contexts in which they occur to reflect etymology. Furthermore, etymology in the sense of 'origin' is often insufficient for predicting letter-sound correspondences. *Chef* and *chief*, for example, are both from the same French root, yet have different mappings for initial *ch*. This results from *chief* being borrowed before initial /tš/ in French was leveled to /š/, and *chef* being borrowed afterwards. Similarly, *machine* and *chorus* are both from Greek, yet also have different correspondences for *ch*.

A deeper issue in identifying features is reflected in the forms *Indian* and *cordial*. Viewed only in relation to *d*, *Indian* appears to have the most expected pronunciation, while *cordial*, *educate*, and a few other words have a less common correspondent. Viewed, however, in relation to the prepalatal consonants that palatalize before a glide or high front vowel (s, z, t), *cordial* contains the predictable mapping and *Indian*, the unpredictable one (cf. issue, pleasure, bastion).

From a linguistic standpoint, there is no simple metric for determining features upon which pronunciation mappings can be made. Therefore, the psychological reality upon which E. J. Gibson originally attempted to base the notion of spellling pattern or upon which Baron and Strawson (1976) and others have attempted to derive *regular* and *irregular* pronunciations remains to be determined.

Consider, as a further example, the vowel mappings in *con* and *cone*. They are both predictable from their graphemic environments and therefore tend to be classed as *regular*. However, exceptions exist; e.g., *ton, glove*. How many such exceptions are required to invalidate the *regular* classification? Compare this example to that of *g* which, like *c*, tends to have one mapping (/ǰ/) before mid and high front vowels (e.g., giraffe, gem) and another (/g/) elsewhere. But many exceptions to this pattern exist (e.g., girl, get, give, gift, gear), and the majority of these exceptions are high frequency words. From a linguistic standpoint, we would probably establish the *g* and *c* patterns in parallel, and treat *girl, give*, etc., as exceptions. Studies of letter-sound learning (Venezky, 1976) show, in contrast, that while readers acquire the *c* pattern by the end of elementary school, *g* continues even with college students to be treated as having a single, invariant mapping /g/. Thus, while letter-sound translation may play a role in word recognition, the concepts of *regular* and *irregular* cannot be completely adduced from linguistic evidence alone.

Orthographic Structure

A second type of orthographic information that has been incorporated into word recognition models has centered not on regularity of spelling to sound mapping, but in the regularity of the letter sequences themselves relative to what occurs most commonly in English words (e.g., Mason, 1975; Massaro, Venezky, and Taylor, 1979). Two types of structure or redundancy have been identified: statistical and rule-governed. *Statistical redundancy* applies to those measures that can be derived solely from the frequency with which letters, letter sequences, and words occur in natural texts. Such measures might be based upon word types or word tokens, and might apply to letters, letter sequences, or to any classification of these entities (e.g., vowel, consonant, spirant). *Rule-governed regularity* applies to those properties of letters and letter sequences which can be generalized from either the phonological constraints of English words or the scribal conventions which control the selection of letters for either representing sounds or marking graphemic, graphotactic, or morphological functions. Rule-governed regularity is usually considered to be dichotomous, although through the specification of rules for generating legal pseudowords, levels of illegality can be defined.

Statistical Redundancy

Ordered approximations to English, based on algorithms proposed by Shannon (1948), were the first descriptions of orthographic structure utilized by psychologists. Miller, Bruner, and Postman (1954) used a random selection technique suggested by Shannon (1948) to generate eight-letter pseudowords for a full-report task. Their fourth-order approximations included items like *mossiant, oneticul, preveral,* and *favorial,* all quite English-like. Wallach (1963), on the other hand, truncated two letters from each string used by Miller et al. thus occasionally producing a non-English beginning or ending (e.g., *mossia, loriai*). Although Miller et al. (1954) were primarily concerned with information load, later experimenters (e.g., Lefton, Spragins, and Byrnes, 1973; Mewhort and Beal, 1977) have advocated ordered approximations as psychologically relevant descriptions of orthographic structure.

A major shortcoming of ordered approximation descriptions is their inability to control structural variations within approximation levels. Although any large group of stimuli generated by this technique will approximate the frequencies of occurrence of English letter sequences, there is no metric incorporated into the method for assessing the relative degree of orthographic structure of strings within order classes.

A second shortcoming of a strict psychological application of this approach is its total reliance on left-to-right generation of letters during word recognition. For example, given bl _ _ _ and bl _ st as partially resolved strings, the probabilities of identifying the third letter in each on partial (or no) featural information would be the same in that information to the left but not to the right of any letter position is relevant for processing.

To overcome the first difficulty, summed bigram and trigram counts have been proposed as metrics. Underwood and Schulz (1960) tabulated the bigrams and trigrams in 2,090 words sampled from the Thorndike and Lorge (1944) word list. Since each sample word was weighted by its frequency of occurrence in texts, token-based bigram and trigram counts resulted. By summing the frequencies of all the bigrams or trigrams within a letter string, a familiarity index can be derived. Such a procedure was used by Anisfeld (1964), for example, in reanalyzing the Gibson, Pick, Osser, and Hammond (1962) results.

The Underwood and Schulz (1960) tables provide a means for differentiating familiar and unfamiliar strings within approximation classes, but suffer from lack of attention to word length or position of letter string within a word. A common suffix like *ies* would heavily weight a trigram count for a pseudoword whether in initial or final position according to these tables, yet in English words, this sequence never occurs initially. Although strong evidence is lacking for demonstrating letter position within a word as a psychologically relevant variable, the Underwood and Schulz (1960) tables could not be used for exploring this factor.

This shortcoming has been overcome in Mayzner and Tresselt (1965) and

Mayzner, Tresselt, and Wolin (1965a, 1965b), which give bigram, trigram, and tetragram counts by word position and by word length for words from three to seven characters in length. Mayzner and Tresselt (1965) also give single-letter positional frequencies, which have been used by Mason (1975) to explore differences between good and poor readers in a target search task. Although Mason (1975) confounds rule-governed orthographic structure with positional frequency, a small but significant positional frequency effect remains when these variables are separated (Massaro, Venezky, and Taylor (1979)).

The relationships among single-letter positional frequency, bigram positional frequency, and trigram positional frequency have yet to be explored. Experimenters who use single-letter or bigram frequency tend not to control for differences in the next higher order unit, thus confounding cannot be ruled out. It is possible, however, to generate pseudowords while independently varying summed single-letter positional frequency and summed bigram positional frequency counts.

With some straining, the same can be done for summed bigram and summed trigram counts. For example, *nachim* has (according to Mayzner, Tresselt, and Wolin, 1965a, 1965b) legal bigrams but no legal trigrams. *Nather,* on the other hand, has legal bigrams and legal trigrams. A legal bigram or a legal trigram means that the spelling pattern occurs with frequencies above zero in the Mayzner, Tresselt and Wolin tables. (The Mayzner tables, although adequate for single letter and bigram counts, are based on too small a sample for reliable trigram counts.)

A further issue that needs to be resolved is the relative importance of token measures and type measures. Token counts represent a summed occurrence for some entity, regardless of the number of unique words which contribute to this sum. Type counts, on the other hand, represent a sum of the unique words in which an entity occurs, regardless of the frequencies of occurrence of the words. That the two are not equivalent can be seen from the tabulation of summed token and type counts in Table 1, based on counts generated from Kucera and Francis (1967).

TABLE 1
Summed Token and Type Counts

Pseudo Word	Summed Counts Tokens	Summed Counts Types
STAP	722	40
TOLT	3271	55
WIFY	11340	26
THET	34517	31

In all of the experimental work reported so far that uses sequential redundancy measures, token counts have been employed. (Solso and King (1976) report type-based tabulations for bigrams, based on the Kucera and Francis (1967) word lists, but no experimental data.) However, there is no *a priori* reason for assuming that word tokens have greater psychological reality than word types, nor are there empirical data relevant to this question.

Type-based measures could be derived from a representation of an internal visual lexicon, but token measures also require the existence of distributional information in memory. While it is reasonable to assume that some forms of both token and type information have psychological reality, it has been argued that the direct application of raw token measures in word recognition is implausible (cf. Savin, 1963). However, some interesting approaches to using token measures embedded within type measures have been proposed by Solomon and Postman (1952), Savin (1963), and more recently by Landauer and Streeter (1973).

But whatever application might be shown for frequency information in word recognition, the specific role of these measures would remain to be solved. If summed frequency (token or type) is psychologically relevant, then the utilization of a measure like bigram frequency would require a maximizing process. Consider the task of resolving a letter string where the visual information allows an acceptable resolution of all letters except one, e.g., $L_1 L_2 - L_4 L_5$. Assume that no visual information were available for L_3. The use of bigram information requires that L_3 be selected to maximize the sum of the bigram frequencies of $L_2 L_3$ and $L_3 L_4$, selecting L_3 from a pool of all 26 letters. If some visual information were available for L_3, then the pool of available letters might be restricted to those that were compatible with the available information but otherwise the strategy would be the same. For a trigram strategy, a similar process could be argued, but with maximizing of the combined frequencies for $L_1 L_2 L_3 + L_2 L_3 L_4 + L_3 L_4 L_5$. The complexity of this strategy makes its psychological relevance suspect.

A further drawback to summed single-letter, bigram, and trigram positional counts, particularly where token counts are utilized, is the dominating role which a small group of high frequency words play. For the four-letter words, as an example, the 37 most frequent words of this length account for approximately one-half of all tokens of four-letter words. This guarantees, on the average, a high, positive correlation between different statistical measures. However, low summed single-letter positional frequency strings will usually be low in summed positional bigram and trigram frequency, but strings high in summed single-letter positional frequency might not be high in the other two measures.

Furthermore, since the extremely high frequency words in English are mostly three, four, and (to a lesser degree) five letters in length, the influence of the highest frequency words necessarily decreases with increasing word length. Thus, we would expect a lower correlation between bigram and trigram counts in seven-letter strings than in four-letter strings.

RULE-GOVERNED REGULARITY

Phonologically-Based Descriptions

Sequential schemes, even when position and word lengths are observed, can generate pseudowords which violate English phonological and spelling conventions. For example, *thrsm, sthse,* and *whrst* are unpronounceable psuedowords which have relatively high summed, position-sensitive bigram counts. These can be contrasted with *bipon, slevy,* and *dufip,* which have relatively low summed bigram counts, but are legal from a rule-governed standpoint. Violations can be avoided, for the most part, by using position sensitive, fourth-order approximations which are also high in summed trigram counts. But this seems unduly burdensome. An alternative is to utilize pronounceability, as described above. Yet for classifying pseudowords, pronounceability alone is not sufficient in that English orthography is constrained by both phonological and scribal conventions.

Phonological Constraints

The most important constraint on English orthography derives from the allowable sequences of sounds in English words. Thus, /sk/ is an allowable initial consonant cluster and therefore spellings for it exist (e.g., sk, sc, sch); on the other hand, /tl/ and /dl/ do not occur and never did occur initially in English words and therefore no spellings for these sequences occur in English words (but they do in Tlingit and several other languages in which initial /tl/ and /dl/ are legal). Complicating this relationship between sound and spelling, however, are several factors, including scribal pedantry, which leave us with spellings for (1) sound sequences which once occurred, but have since been dropped from the language (e.g., wr-, kn-, gn-), or for (2) sound sequences that were presumed to have occurred in the language from which a word was borrowed (e.g., *psychology, ptarmigan, debt*). For *psychology,* this assumption is supportable; for *ptarmigan,* it is clearly based on mistaken etymology; and for *debt* (and various others), it is only partially true, in that *b* occurred in the Latin ancestor of *debt (debitum),* but not in a final cluster. (The *b* in English *debt* was added in the late fifteenth century to reflect the word's classical origin.)

A further complication is the deviation in serial order of several spellings from the order of sounds which they represent. Chief among these is *wh,* a spelling which at least some English speakers have retained as /hw/, and -le as in bottle, etc. which spells (in deliberate speech) /əl/. (*wh-* and *-le* are scribal reversals of earlier *hw-* and *-el.*)

But even with these deviations from expected practice, most English sounds are represented in a moderately rational fashion in the orthography (Venezky, 1970). However, since English possesses more than one potential spelling for almost every sound, a second set of constraints, here called scribal constraints, enter into rule-governed orthographic structure.

Scribal Constraints

Unlike Spanish, English resists doubled letters in word-initial position. Only a handful of exceptions to this constraint exists, primarily involving the form *oo* (e.g., *oodles, ooze, aardvark*). In other positions, only 15 letters can double; those that don't (or rarely do) are *a, h, i, j, k, q, u, v, w, x, y.* But there are further restrictions. With few exceptions, doubled consonants do not occur after secondary vowels, (e.g., ai, ou, ey) and in final position only *ff, ll,* and *ss* are common. Finally, the digraph units are not allowed to double. For one, *ch,* a pseudogeminate *(tch)* was adopted in Early Modern English, at about the same time that *dg* was adopted to represent doubled *g* when it stood for /ǰ/, and *ck* was invented for doubled *k,* which now occurs only in inflected forms of *trek* (e.g., trekking).

A second set of graphemic constraints regulate the distribution of *u—v—w—* and *i—y.* From the Middle English period up to about the middle of the seventeenth century, *u* and *v* were used indiscriminately to represent /u/ and /v/, although *v* tended to be used initially and *u* elsewhere. Digraphs like *ou,* when written before a vowel might be mistaken for *ov;* therefore *u* was doubled, and eventually written *w.* In time, *u* in final position as the second part of a secondary vowel spelling was replaced by *w,* giving the present-day alternations *au/aw, eu/ew,* and *ou/ow.* In general, the *w* variant occurs before vowels and in final position; the *u* variant occurs elsewhere. Where *u* would occur as a primary vowel spelling in final position, a final *e* is added (e.g., blue).

For reasons not altogether clear, but related to the ambiguity of *u-v* spellings, *v* was also exorcised from final position. But since no simple replacement was available as was for *u*, an *e* was added after the letter, thus giving present-day spellings like *dove, love,* and *have,* where the final *e* does not mark a free vowel pronunciation.

Similarly i-y now alternate exactly as u and w do, as second elements of digraph vowels Thus, *maid* but *may; oi* but *boy;* and *either* but *grey.*

A third set of constraints apply to short words. Since at least the time of Noah Webster and probably before, English lexicographers have avoided extending the repertoire of two-letter words beyond a small group of function words. Two mechanisms were adopted for lengthening two-letter words that didn't qualify for inclusion in this group: doubling the final consonant, as in *ebb, add, odd, egg,* and *inn,* or adding a final *e,* (e.g., *doe hoe, dve,* and *lye*). The former mechanism has created unique final consonant clusters in that *b, d, g,* and *n* do not double in final position in any other English words.

Since no English language academy ever imposed a grand design on English spelling, what general principles there are result from the confluence of various scribal traditions with the periodic ravages of sound change unaccompanied by spelling change. Yet a principle of sorts does emerge from the current system; viz, "Accede to all visual interests up to, but not beyond the point where passage from spelling to sound is impaired."

Psychological Assumptions

Rule-governed descriptions of orthographic structure differ from other definitions of orthographic structure in two aspects. First, in the utilization of word types rather than word tokens, and second, in classing as regular certain spellings that do not occur (yet) in English words and classing as irregular some that do. The first difference is not essential to the definition of orthographic structure, but it is probably more a reflection of the differing epistemologies of linguistics and psychologists. The second difference, however, is central to the concept of orthographic structure and requires further explication.

The primary assumption behind the orthographic rules presented here is that readers abstract from their experiences with printed words those spellings that are productive in present-day English, and use these to facilitate recognition. For example, silent initial letters, regardless of the frequency of occurrence of the words in which they occur, are classed as irregular since it is unlikely that new words brought into the English language would ever be spelled with them. (Spelling analogies might present an exception to this assumption and cannot be ignored. However, it is unlikely that silent initial spellings would ever be widely extended.)

On the other hand, the initial sequences *voa-* and *woa-* and the final sequences *-eng, -erl,* and *-ofe* do not occur in the first 20,000 words in the Thorndike and Lorge (1944) list, yet they are classed as orthographically regular since they do not violate any phonological or scribal pattern. Thus, pseudowords like *voam, woach, treng, derl,* and *strofe* are pronounceable and structurally regular, even though they contain trigrams which do not occur in English words. The recognizability of such strings relative to similar pseudowords with higher summed trigram counts is an issue that remains to be investigated.

Dichotomy Issue

A second issue which merits empirical investigation concerns the assumed dichotomous nature of orthographic regularity. Given the different constraints on orthographic structure discussed above, we should question whether or not the visual and phonological elements are separated psychologically. Are, for example, irregular strings like *clav* and *cklib* processed differently from an illegal string like *tprif?* All are illegal and non-occurring, yet they differ in type of illegality. *tprif* contains an initial consonant cluster than is phonologically illegal and therefore scribally non-occurring. *Clav* and *ckib* are both phonologically acceptable, but contain scribally outlawed features, viz, final *v* in *clav* and initial *ck* in *cklib.* In, for example, a lexical decision task, do subjects reject phonologically illegal strings before visually illegal ones? If so, must we then postulate two different mechanisms for the lexical decision task: a phonological mediation process, á la Rubenstein, Lewis, and Rubenstein (1971), and then a structural screening? A partial answer to these questions has been provided by Hart (1980) who has found differential responses in subjective ratings of 'wordness' and in a lexical decision task to different types of structural violations.

Finally, the processes through which orthographic regularity facilitates recognition in different experimental tasks need to be examined. Consider, for example, the orthographically irregular pseudoword *kclum.* Does the probability of correctly identifying a target letter depend upon whether it is in the initial portion of the string, which is irregular, or the final portion, which is regular? In other words, we might find no differences between orthographically regular and orthographically irregular strings if we restrict our recognition task to operations on the regular substrings.

THE FALLACY OF PHONOLOGICAL MEDIATION

Current models for printed word recognition tend to fall into two major classes: mediated and unmediated, where mediation is generally assumed to be phonological. Representative of the mediated class are the models proposed by Gough (1972), Rubenstein, Lewis and Rubenstein (1971) and Coltheart et al. 1977). Although some evidence against the necessity for phonological mediation in word recognition has been offered (e.g., Kleiman, 1975; Levy, 1978; Massaro, 1977; Tzeng, Hung, and Wang, 1977), the appeal of this model is still quite high. It may be informative, therefore, to inspect both logically and empirically the rationale of phonologically mediated word recognition, particularly in relation to the lexical decision paradigm which has provided the primary support for this position.

The Logical Basis of Phonological Mediation

Phonological mediation, as expressed in the experimental literature of the past two decades, is not a well-defined process for recognition of printed words. Instead, it tends to be an inchoate concept of recognition, centered on searching a phonological lexicon that has been constructed over time in long term memory. Although few attempts have been made to explicate the tasks involved in such a procedure, it is possible to define the range of possible stages that would be required for phonological medication (P.M.) to be consistent with the empirical data on printed word recogition. As will be shown below, it is also possible, through appeal to common sense, to narrow this range of possible procedures.

In its purest form (cf. Gough 1972), P.M. requires that a fully resolved string of letters be recoded into a phonological form, and this form then be matched against entries in a listening lexicon. Recoding from visual to phonological forms is assumed to be done via letter-sound rules which have been abstracted from prior reading experience. It seems reasonable to assume that actual rules, in some abstracted form, are utilized, rather than an organized search of the lexicon for words with similar letter patterns. If the latter could be done for letters in their various intraword positions, including medial and final, then phonological recoding would be unnecessary. That is, if a procedure were available for rapid location and matching of letter strings within words, then it is reasonable to assume that this same system could match the entire strings also. Therefore, we must assume that generalized letter-sound relationships of some form are utilized. On the nature of such generalizations, more will be said below.

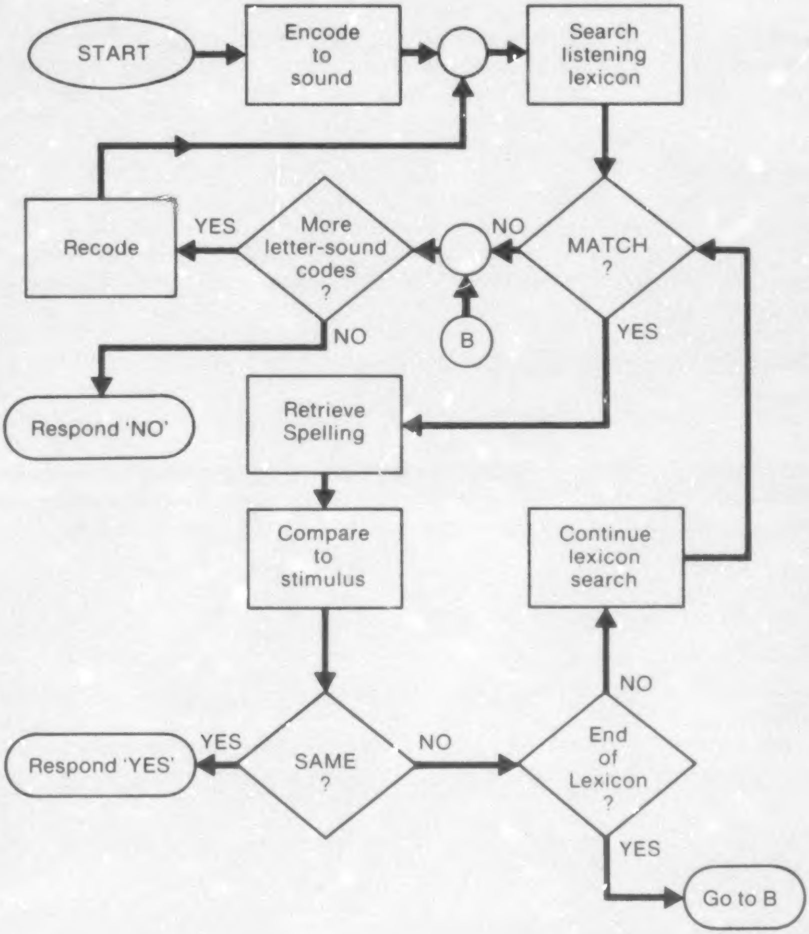
At this point I am assuming that the means by which the visual input is resolved is unimportant for testing the P.M. hypothesis. (There may be a slight advantage to the left-to-right serial resolution model, as advocated by Gough (1972), especially when incorporated with a search with partial phonological information as suggested by Rubenstein, Lewis, and Rubenstein (1971), but there is limited empirical support for such a position.) Furthermore, I am avoiding speculation at this point on the relative times required by the various stages. Notice, however, that the steps suggested above must be supplemented with several additional stages to account for the homophone effect that has been found in a number of lexical decision studies (e.g., Coltheart et al. 1977). In brief, pseudowords that are generally pronounced identically to real English words take longer to reject in the lexical decision task than pseudowords that are not. Although this effect is considerably more complex than is assumed in most studies, the only aspects of it that are important for now are (a) the effect has been replicated on a number of occasions, and (b) error rates on pseudohomophones are typically low (2-8% range).

Although I will offer later a different hypothesis for the origin of the homophone effect, the low error rate is what is of immediate interest. This implies that if a match is made between a pseudohomophone and an entry in the listening lexicon, the subject does not simply respond 'yes'. Instead, a visual match must be made between the input string and the spelling for the matched entry in the listening lexicon. To account for this low error rate, we must assume that first, spellings are readily accessible from phonological entries in the lexicon, and second that the subject has the ability to compare such spellings to the input strings.

These assumptions make the model sketched above consistent with the empirical data for certain classes of pseudowords, but this model does not yet handle real words that have irregular letter-sound correspondences. Consider what happens, for example, when the first phonological recoding of a real word does not produce the correct pronunciation for that word. If the incorrect pronunciation does not match any entry in the listening lexicon, the subject cannot respond 'no'; if he did, he would have an abnormally high error rate on such common words as *who, whole, glove,* and *was.*

Consequently, for the P.M. model to work, a 'no' response must not be made on the first lexicon search failure for a stimulus string. Instead, a new pronunciation has to be assigned to the string, and the search repeated. How often to recycle on search failure is a difficult problem to answer. Consider, for example, a short (and highly frequent) word like *whom.* If pronunciations are initially assigned based on the most frequent correspondences for the letters involved, then the first pronunciation generated for this string would be /hwam/, which will not match any entry in the listening lexicon of a Midwestern English speaker. How the pronunciation would be altered for recycling is a matter of speculation. If we work left to right, and from most frequent to least frequent correspondences, the initial sequence of phonological strings to be tested after /hwam/ would be /ham/, /hwom/, /hom/, /hw/m/, and /h/m/. Exactly when (or how) the pronunciation /hum/ would be assigned is difficult to predict, given that *whom* is one of a small number of words in English where *o* is pronounced /u/. But assuming that /hum/ would eventually be assigned, the model shown in Figure 1 would apply.

Figure 1. A Processing Model for Phonological Mediation


START
Encode to sound
Search listening lexicon
Recode
YES
More letter-sound codes ?
NO
MATCH ?
NO
B
YES
Respond 'NO'
Retrieve Spelling
Compare to stimulus
Continue lexicon search
NO
Respond 'YES'
YES
SAME ?
NO
End of Lexicon ?
YES
Go to B

At this point it is informative to consider several alternatives to Figure 1. On one hand, the search system might not require an exact phonological match. It could, for example, test the spelling of every phonological string that was within a specified distance of the recoded input, where 'distance' was defined by some as yet unspecified phonological metric. If this were practicable, then we might eliminate the need for much of the recycling. It is not immediately obvious how such a metric would be derived, but it is clear that any approximation algorithm would lengthen reject times for non-words as well as shorten acceptance times for real words.

Furthermore, for multisyllabic pseudowords, enormously large numbers of words would need to be tested. Consider, for example a pseudoword with eight phonemes, three of which are vowels (e.g. plesipom: /plízipàm/). If each phoneme, on the average, were assumed to have two phonological neighbors that should be tested, then the total number of words to test would be 3^8 or 6,561. Even if we assume that only the vowels would be allowed to vary, the number of alternatives would seem to be beyond what might be tested in the 600-900 milliseconds that most studies report for rejection times. For these reasons, the model proposed in Figure 1 will be used in the next section as a basis for empirical tests of the phonological mediation hypothesis.

Empirical Assessment of Phonological Mediation

The studies summarized below were all performed with subjects drawn from the undergraduate student body at the University of Delaware. A lexical decision paradigm was used in each study, with a letter string of five or six letters in length exposed for up to 2.5 sec. Subjects responded by pressing telegraph keys; *yes* with their dominant hand for a real word and *no* with their nondominant hand for a pseudoword. Response time was measured from the exposure of the stimulus until depression of the response key. Stimuli were presented on a Tektronix 604 viewscope in a lower case font developed at the University of Wisconsin (Massaro, et al., 1980).

Stimuli presentation and response recording were controlled by a PDP 11/03 computer. Instructions to the subjects emphasized both speed and accuracy, with primary emphasis on accuracy. The presentation sequence began with 20 trials in which the subject responded to the words *yes* and *no*. This was followed by 96 warm up trials with words and pseudowords similar to those in the

TABLE 2
Real Word Sample for Experiment 1

Low Frequency		High Frequency	
Irregular	**Regular**	**Irregular**	**Regular**
ache	auto	been	baby
bull	chop	done	flat
chef	heel	hour	hope
pint	omen	sure	onto
apron	ample	above	angry
comet	basin	dozen	birth
gauge	cliff	often	later
steak	scalp	woman	teeth

experimental trials, and then 240 experimental trials, presented in randomized blocks of 60 times each.

Experiment 1

In experiment 1, real words were selected to give two levels of word frequency (high and low), based on Kucera and Francis (1967), and two levels of letter-sound regularity (high and low). Representative stimuli are shown in Table 2.

Pseudowords were generated for the following four classes:

1. Primary homophones: words whose most common pronunciations yielded homophones of English words (e.g., bild, mone, coave, trane).
2. Secondary homophones: words whose most common pronunciations would not yield homophones, but whose alternate pronunciations would (e.g., cheg (=keg), sise (=size), fleak (=fleck), swort (=sort)).
3. Structurally regular, non-homophones (e.g., drin, rish, hamit, slaft).
4. Structurally irregular but pronounceable non-homophones (e.g., ckif, trav, hwoty, tyroi).

The real words were designed to contrast a letter-sound effect with a frequency effect. If phonological mediation occurred, then words with regular letter-sound correspondences should be accepted faster than words with irregular correspondences. But since letter-sound regularity might interact with word frequency, these two factors were crossed in the design.

The psuedowords provided an additional strong test for phonological mediation. By the argument given above, pseudowords whose primary (or most common) pronunciations yield homophones should not differ significantly in rejection times from pseudowords whose primary pronunciations do not yield homophones, but whose secondary pronunciations do. Consider, for example, the processing of the secondary homophone *cheg*. The primary pronunciation of this word is /čeg/, which does not match any entry in the listening lexicon. But the subject cannot respond 'no' at this point; if he or she did, abnormally (and consistently) high error rates would occur for *chord, chorus,* and various other forms, in that their primary pronunciations would also yield non-words.

Therefore, for phonological mediation to function smoothly, alternative pronunciations would need to be generated and tested against the listening lexicon. At some point *cheg* will be assigned the pronunciation /kɛg) which will match the representation for *keg* in the listening lexicon. Therefore, whatever delay in responding occurs from the *build* representation of *bild* (a primary pseudohomophone) must also occur from the *keg* representation of *cheg*.

The remaining two pseudoword groups were included as controls. According to the phonological mediation hypothesis, groups 1 and 2 should take significantly longer to reject than groups 3 and 4, while groups 3 and 4 should not differ significantly in rejection times. If, on the other hand, the lexical decision task incorporated purely visual matching, group 4 should be rejected most rapidly since it contained the most irregular forms, while groups 1, 2, and 3 should not differ in rejection times.

Results and Discussion

The median response latencies, shown in Table 3, were entered in separate ANOVAs for real and for pseudo words.

TABLE 3
Median Response Latencies in Milliseconds
(and error percentages)
Experiment 1 (N=27)

	Real Words	
	High Frequency	Low Frequency
Regular	649 (1)	733 (6)
Irregular	650 (2)	740 (7)

Pseudowords			
Primary homophones	**Secondary homophones**	**Regular non-hom.**	**Structurally Irregular**
916 (7)	886 (4)	864 (4)	744 (1)

For real words a 2 × 2 analysis of variance (regularity × frequency) showed a huge effect for word frequency ($F=135.1$, $p<.005$), but no effect for letter-sound regularity ($F<1$), and no frequency-regularity interaction ($F<1$).

For pseudowords, a one-way analysis of variance showed a large pseudoword type effect ($F=58.2$, $p<.005$). A priori t-tests, however, revealed that while primary homophones were significantly different from non-homophones, secondary homophones were not. Furthermore, primary and secondary homophones were significantly different ($t=30$, $p<.05$). Structurally irregular pseudowords were rejected significantly faster than all other types of pseudowords.

Thus, the phonological mediation view was clearly rejected in the real word analysis, but given a mixed review in the pseudoword analysis. Response times to primary and secondary homophones did differ significantly, although the difference was hardly overwhelming. Furthermore, structurally irregular pseudowords were rejected significantly faster than each of the other types, a result which phonological mediation cannot account for. On the other hand, primary homophones clearly required more time to reject than non-homophones. This result made further testing necessary.

Experiment 2

The real words in experiment two were nearly identical to those in experiment 1 and were included as a replication of the experiment 1 real word results. The pseudowords, however, incorporated two modifications. Secondary homophones were replaced by structurally irregular homophones (e.g., *brij, ckuf, bbaik, glovv*), and the two non-homophone lists were replaced by structurally irregular lists in which initial position irregularities (e.g., *aarp, kwig*) were contrasted with final position irregularities (e.g., *tecc, parnu, wealg*).

Results and Discussion

The median response latencies shown in Table 4 were submitted to separate ANOVAs for real and for pseudowords.

TABLE 4
Median Response Latencies in Milliseconds
(and error percentages)
Experiment 2 (N = 18)

	Real Words	
	High Frequency	Low Frequency
Regular	658 (2)	744 (3)
Irregular	659 (2)	739 (3)

Pseudowords			
Regular homophones	Irregular homophones	Final irregular	Initial irregular
980 (6)	723 (1)	717 (1)	717 (1)

As in experiment 1, a large frequency effect ($F = 33.3$, $p < .005$), but no regularity effect ($F < 1$) or interaction ($F < 1$) occurred for the real words. For the pseudowords, a large class effect occurred, as expected ($F = 60.6$, $p < .005$), but a priori *t*-tests showed sigificant differences in response times only for regular and irregular primary homophones. The irregular, primary homophones did not differ significantly from either of the non-homophone classes, nor did initial and final irregularities differ.

Thus, a further complication for the phonological mediation hypothesis was added, in that structurally irregular homophones failed to show a homophone effect. This can be accounted for only by a visual screening process whereby structurally irregular strings are rejected visually before they are recoded to sound. This does not eliminate the possibility of recoding for those strings that pass the screening, but it does eliminate the possibility of phonological recoding prior to a spelling test.

Alternative Hypothesis

These two studies, plus a third not completely analyzed yet, indicate several weaknesses in the phonological mediation hypothesis. Yet the persistence of the homophone effect in Experiment 1 requires a processing model that includes phonological recoding at some stage. Shulman, Hornak, and Sanders (1978), using a version of the lexical decision task that differs from the one employed here, have also found conditions under which phonological recoding appears not to occur. They hypothesize that the phonetic code, when it is generated, is used to retain the stimulus string in short term memory (STM) while an exhaustive (visual) search is made. Thus, even when recoding occurs, it does not mediate lexical access (cf. Coltheart et al., 1977).

The results found here strongly support this view. If we assume that *all* strings are recoded phonologically for retention in STM while visual search proceeds, then we can attribute the primary homophone effect to interference. That is, while the visual search is reaching a 'no' conclusion, the phonological form inadvertently makes contact with its homophonic representation in long term memory. The resulting conflict between 'matched' and 'not matched' indicators then requires extra time to resolve.

The key to fitting this model to the data is the assumption that the simplest phonological representation for a string is generated, strictly as an aid to STM retention and not as a mediator for recognition. Therefore, no phonological recycling occurs and the homophonic nature of secondary homophones is (in general) not discovered. This hypothesis suggests the existence of a basic set of letter-sound correspondences that are invoked for STM recoding. This set probably has very limited representation for the more complex variant patterns; nevertheless, some variant patterns can probably be stored (e.g., the final e-final consonant patterns for vowels).

All of the foregoing has been derived for the lexical decision task which, though reading-like, is clearly not reading. What can we assume about normal reading from all of this? First, the lack of even a hint of a letter-sound regularity effect with real words leaves phonological mediation in the same class as the chimera the unicorn, and the succubus. On the other hand, the robust frequency effect ensures that statistical regularity measures will show a strong correlation with response times, even though these relationships might be strictly artifactual. This leaves open the possibility that orthographic factors, whether phonologically related or structural, play a minimal role in the word recognition habits of experienced readers. Structural features might, as suggested by Massaro (1977), facilitate primary recognition of letters, but this probably accounts for little of the variance in total recognition time.

Apparent demonstrations of orthographic effects in word recognition (e.g., Baron and Strawson, 1977) rarely reappear under careful replication (cf. Mason, 1978). Thus, we are left with a curious phenomenon. Orthography plays a critical role in learning to read, and in a variety of reading-like experimental tasks (e.g., letter search, lexical decision); however, strong logical and empirical evidence for a major role of orthography in word recognition is lacking. In time we may gain a detailed understanding of human information processing for the lexical decision task and for various related reading-like tasks. It is doubtful, however, that any stronger relationships between these tasks and the normal reading process will be discovered than exist now.

REFERENCES

Anisfeld, M. A. Comment on "The role of grapheme-phoneme correspondence in the perception of words." *American Journal of Psychology, 1964, 77,* 320-326.

Baron, J. M. & Strawson, C. Use of orthographic and word specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, *2,* 386-393.

Chomsky, N., & Halle, M. *The sound pattern of English.* New York: Harper & Row, 1968.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI.* Hillsdale, NJ: Lawrence Erlbaum, 1977.

Gibson, E. J., Pick, A., Osser, H., & Hammond, M. The role of grapheme-phoneme correspondence in the perception of words. *American Journal of Psychology,* 1962, *75,* 554-570.

Gibson, E. J., Shurcliff, A., & Yonas, A. Utilization of spelling patterns by deaf and hearing subjects. In H. Levin and J. P. Williams (Eds.), *Basic studies on reading.* New York: Basic Books, 1970.

Gough, P. B. One second of reading. In J. F. Kavangah and I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, MA: MIT Press, 1972.

Hart, J. *The psychological reality of different spelling patterns.* Unpublished doctoral dissertation, University of Delaware, 1980.

Kleiman, G. M. Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior,* 1975, *14,* 323-339.

Kucera, H., & Francis, N. W. *Computational analysis of present-day American English.* Providence: Brown University Press, 1967.

Landauer, T. K., & Streeter, L.A. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior,* 1973, *12,* 119-131.

Lefton, L. A., Spragins, A. B. & Brynes, J. English orthography: relation to reading experience. *Bulletin of the Psychonomic Society,* 1973, *2,* 281-282.

Levy, B. A. Speech analysis during sentence processing: Reading and listening. *Visible Language,* 1978, *12,* 81-101.

Lounsbury, T. R. *English spelling and spelling reform.* New York: Harper & Brothers Publishers, 1909.

Mason, M. Reading ability and letter search time: Effects of orthographic structure defined by single-letter positional frequency. *Journal of Experimental Psychology,* 1975, *104,* 146-166.

Mason, M. From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. *Memory and Cognition,* 1978, *6,* 568-581.

Massaro, D. W. *Reading and listening* (Technical Report No. 423). Madison: Wisconsin Research and Development Center for Individualized Schooling, December 1977.

Massaro, D. W., Taylor, G. A., Venezky, R. L., Jastrzembski, J. E., & Lucas, P. A. *Letter and word perception.* Amsterdam: North-Holland, 1980.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General,* 1979, *108,* 107-124.

Mayzner, M. S., & Tresselt, M. E. Tables of single-letter and bigram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements,* 1965, *1* (No. 2).

Mayzner, M. S., Tresselt, M. E., & Wolin, B. R. Tables of tetragram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements,* 1965,*1,* (4, Whole No. 4). (a)

Mayzner, M. S., Tresselt, M. E., & Wolin, B. R. Tables of trigram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements,* 1965, *1* (3, Whole No. 3). (b)

Mewhort, D. J. K., & Beal, A. L. Mechanisms of word identification. *Journal of Experimental Psychology: Human Perception and Performance,* 1977, *3,* 629-640.

Miller, G. A., Bruner, J. S., & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology,* 1954, *50,* 129-139.

Pitman, I. J. Learning to read: An experiment. *Journal of Royal Society of Arts,* 1961, *109,* 149-180.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior,* 1971, *10, 645-657.*

Savin, H. B. Word-frequency effects and errors in the perception of speech. *Journal of the Acoustical Society of America,* 1963, *35,* 200-206.

Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal,* 1948, *27,* 379-423; 622-656.

Shulman, H. G., Hornak, R., & Sanders, E. The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory and Cognition,* 1978, *6,* 115-123.

Smith, E. E., & Spoehr, K. T. The perception of printed English: A theoretical perspective. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, NJ: Lawrence Erlbaum, 1974.

Solomon, R. I., & Postman, L. Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology,* 1952, *43,* 195-201.

Solso, R. L., & King, J. F. Frequency and versatility of letters in the English language. *Behavior Research Methods and Instrumentation,* 1976, *8,* 283-286.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers College, 1944.

Tzeng, O. J. L., Hung, D. L., & Wang, S. Y. Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, *3,* 621-630.

Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning.* New York: Lippincott, 1960.

Venezky, R. L. English orthography: Its graphical structure and its relation to sound. *Reading Research Quarterly,* 1967, *2,* 75-106.

Venezky, R. L. *The structure of English orthography.* The Hague: Mouton and Co., 1970.

Venezky, R. L. *Theoretical and experimental base for teaching reading.* The Hague: 1976.

Wallach, M. A. Perceptual recognition of approximations to English in relation to spelling achievement. *Journal of Educational Psychology,* 1963, *54,* 57-62.

CONNIE JUEL
University of Texas at Austin

ROBERT L. SOLSO
University of Idaho

THE ROLE OF ORTHOGRAPHIC REDUNDANCY, VERSATILITY, AND SPELLING-SOUND CORRESPONDENCES IN WORD IDENTIFICATION

The isolation of factors involved in the identification of a word and the meaning it represents is among the most problematic areas in reading research. A few factors have been successfully identified which have afforded theorists some significant insights into the reading process. Representative of these theorists are Cattell (1886) who demonstrated the word apprehension effect; Miller, Bruner and Postman (1954) and Rumelhart and Siple (1974) who showed the effect of orthographic redundancy; Mason (1975) and Massaro, Venezky, and Taylor (1979) who demonstrated the effects of letter-positional redundancy; and Calfee, Venezky, and Chapman (1969) and Shankweiler and Liberman (1972) who demonstrated the influence of phonemic structure.

A study by Johnson (1970) is particularly pertinent to the present study. In his experiment word type counts (the number of *different* words that exemplify a letter or phoneme) were found to be more closely related to elementary school student's pronunciations than were word token frequencies (the total frequency of occurrence of letters or phonemes). These two counts, type-token, are sometimes called versatility (number of different words) and frequency (total occurrences of words). Since Johnson's paper over a decade ago, new word norms have appeared which give a more reliable index of versatility and frequency effects. These new counts can enhance our understanding of letter patterns in words. Specifically, the effect of positional factors (the versatility and frequency) of letter(s) in words can now be determined.

We hypothesize that one part of the word identification process involves the transfer of knowledge about positional letter functions, gained through experience with a number of different words. One means of testing this hypothesis is to create a list of words in which positional letters occur in a large number of different words and contrast word identification processing on that list with a list of words whose positional letters occur in a few number of different words. With the aid of high speed computers it has been possible to accurately describe words in terms of their positional versatility, i.e., ones in which the bigrams appearing in numerous other words in the same positions can be identified. Thus, in terms of the above hypothesis, the factor of transfer of positional knowledge to new words can be effectively measured. Of course, it should be shown that versatility affects word identification when orthographic redundancy (total frequencies of letter patterns), letter-sound correspondences, and word frequencies are controlled.

In the two experiments reported in this paper, we evaluated the effect of versatility (and other relevant factors) on a lexical decision and a meaning access task.

METHOD

Two experimental tasks were employed. Task 1 was a lexical decision task in which subjects saw a word (or non-word) projected on a screen. The dependent measure was RT (Reaction Time) to decide if the visual display was a word. In Task 2 subjects saw a word and two pictures displayed on a screen. One picture depicted the meaning of the word and one picture was unrelated to the word. The dependent measure was RT to press the button corresponding to the correct picture. Stimulus words varied on four dimensions: orthographic redundancy, versatility, decodability of spelling-sound correspondences, and word frequency.

Subjects

The two tasks were presented to two groups of subjects. One group consisted of 24 undergraduate students enrolled in an education course at the University of Texas at Austin. The second group consisted of 24 children enrolled in a summer reading program at the University of Texas at Austin. The children were randomly selected from a larger group of second and fifth graders who scored between 88 and 126 on an abbreviated form of the WISC, until six good and six poor second grade readers and six good and six poor fifth grade readers were obtained. (Mean IQ scores for both the high and low ability readers was 102.) High ability readers were those who scored six months to two years above their grade placement on both the comprehension and word recognition subtests of the Woodcock. Low ability readers scored six months to two years below their grade placement on the same measures.

Materials

Sixty-four four-letter words were selected which varied factorially on all possible combinations of the experimental word factors (high/low versatility; high/low orthographic redundancy; easy/hard decodability, high/low word frequency). Versatility and orthographic redundancy measures were based on the summed positional bigram versatility (SOBIV's) and frequency (SOBIF's) norms of Solso and Juel (1980). The mean summed versatility for high versatile words was 87 per million, with a range from 64 to 145. The mean summed versatility for low versatile words was 49 per million, with a range from 9 to 55. The mean summed frequency for orthographically redundant words was 9,202, with a range of 5,695 to 17,157. The mean summed frequency for nonredundant words was 2,216, with a range from 90 to 5,185. Easy decodable words had vowel and consonant sound patterns which corresponded both to their major sound correspondences as cited in Venezky (1970) and represented commonly taught phonic generalizations (e.g., CV̄CE, CV̆C, V̄V). Hard decodable words had either vowel or consonant patterns that corresponded to the minor sound patterns of Venezky and/or represented exceptions to common phonic generalizations. Frequency was determined by the total word frequency count of Carroll, Davies, and Richman (1971). The mean frequency for high frequency words was 1,183 per about five million, with a range from 123 to 5,386. The mean frequency for low frequency words was 41 per about five million, with a range from 2 to 88. Some examples are: A high frequency, orthographically redundant, easy decodable, versatile word is "feet"; a high frequency, orthographically nonredundant, easy decodable, nonversatile word is "tree"; a low frequency, orthographically redundant, easy decodable, versatile word is "bean"; and a low frequency, orthographically nonredundant, hard decodable, versatile word is "pear."

All words appeared in black, lower case print. In the picture task, the pictures were black schematic drawings. The equivalency of difficulty in deciphering the pictures was controlled. Twelve adults, who were not subjects in the experiment, were shown each slide with the two pictures, but with the target word covered. The experimenter orally named the target word while these subjects pressed the corresponding RT button. Mean scores for high and low frequency words were computed. Only picture pairs with RT's within one standard deviation of the appropriate mean were used.

Apparatus

Visual stimuli were displayed on a back projected screen by means of a carousel projector. Each display was shown by means of a remote control switch which simultaneously initiated a 1/100 second timer. A reaction time device was made which consisted of two keys mounted on a wooden board each of which when depressed stopped the 1/100 second timer.

Procedure

Subjects were tested individually. Each subject sat in front of a table with one hand resting on the RT keyboard. The experimenter randomly chose one of the two conditions.

In the lexical decision task the subjects were instructed to look at the four letters on the screen and decide if they formed a real word or a nonword. A real word was defined as one whose meaning they knew. A nonword was defined as a string of letters which did not mean anything.

In the picture task the subject was instructed to read the word on the screen, choose the picture (1 or 2) which represented its meaning, and press the RT key which corresponded to that picture. Prior to the main experiments in both the lexical and picture tasks, several practice slides were presented until response was automatic. The sequence of slides in both tasks was reordered periodically to offset any order effects.

Design

When the subjects were children, both conditions employed a mixed design in which variables were factorially combined. The between subjects factors were Grade (2,5) and Ability (High or Low). The within subjects word factors were Decodability (Easy or Hard), Frequency (High or Low), Orthographic redundancy (Redundant or Nonredundant), and Versatility (High or Low). There were four replications of each word type (i.e., four easy decodable, high frequency, orthographically redundant, versatile words) in the picture task. Three of these replications for *each* word type were randomly selected for the Word/Nonword task. In addition, an equal number of nonwords were used. Half of the nonwords were orthographically legal (e.g., "dila"); half were orthographically illegal (e.g., rglz). With adult subjects in both tasks the within subjects design was the same, but there were no between subjects factors.

Results and Discussion

Data were analyed by repeated measures analysis of variance. Only correct responses were analyzed. The error rate was approximately seven percent for children and three percent for adults. A significance level of .05 was adopted. Analyses were performed on both the raw data with outliers removed (scores in excess of 3 S.D.'s) and on medians. The results of both analyses yielded similar conclusions. Analysis of the raw data appears in this report. Significant main effects and means appear in Table 1 for the lexical decision task and in Table 2 for the picture task. Means in Tables 1 and 2 represent RT's from the instant the visual display appeared on the screen until the key press. Interactions which reached significance are discussed in the text.

Lexical Decision Task

As shown in Table 1, both Grade and Ability had significant main effects for children with older and or better readers having faster RT's. Other significant main effects for children were Frequency and Orthography. However, as shown in Figure 1, these factors both interacted significantly with Versatility (Frequency X Versatility, $F(1,20) = 5.8$, $p < .05$; Orthography X Versatility, $F(1,20) = 5.0$, $p < .05$). In particular, high frequency, nonversatile and redundant, nonversatile words were more quickly identified as words. Poor readers appeared to do more detailed subunit processing than good readers as evidenced by the greater impact of decodability and redundancy on their Reaction times (Ability X Decodability, $F(1,20) = 8.3$, $p < .01$; Ability X Redundancy, $F(1,20) = 7.4$, $p < .05$). Poor readers in the second grade group had a particularly difficult time making lexical decisions (Ability X Grade, $F(1,20) = 7.3$, $p < .05$).

The lexical decisions of adults showed significant main effects of Frequency and Versatility (Table 1). There was also a significant interaction of Frequency and Versatility ($F(1,23) = 16.1$, $p < .01$; Figure 2) indicating that high frequency, nonversatile words are more quickly processed than other word types.

Results on the lexical decision task suggest an initial stage in word processing in which incoming word features are checked against a corpus of high frequency words. If a match is found then no further subunit processing is required. Within this initial stage, nonversatile words would be distinctive and, therefore, more

TABLE 1

ANOVA's and Means for Significant Main Effects in the Lexical Decision Task

	Children		Adults	
Source	F	Means	F	Means
Grade	4.7*			
2		1621		
5		1195		
Ability	16.5**			
high		1007		
low		1809		
Frequency	9.1**		4.6*	
high		1303		599
low		1513		623
Orthography	17.4**			
redundant		1300		
nonredundant		1516		
Versatility			11.5**	
versatile				624
nonversatile				598

(df = 1,20, p<.05 = 4.35, p<.01 = 8.10) (df = 1,23, p<.05 = 4.28, p<.01 = 7.88)
*p<.05
**p<.01

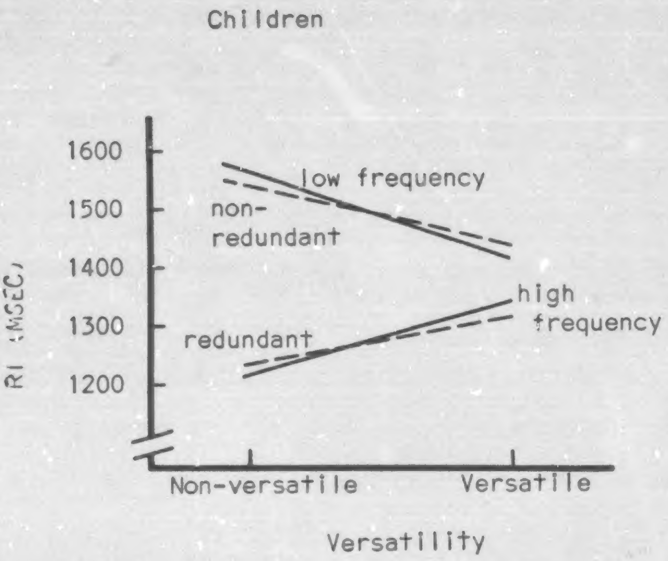


Figure 1: Interactions of Frequency X Versatility and Orthography X Versatility for Children in the Lexical Decision Task.

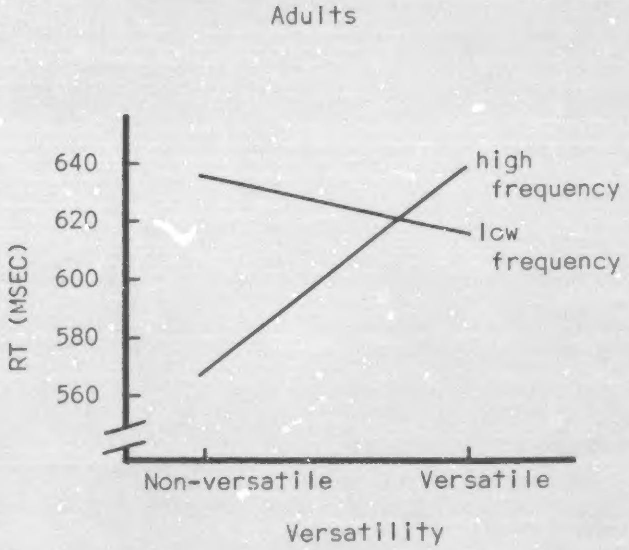


Figure 2: Interaction of Frequency X Versatility for Adults in the Lexical Decision Task.

TABLE 2

ANOVA's and Means for Significant Main Effects in the Picture Task

	Children		Adults	
Source	F	Means	F	Means
Grade	12.5**			
2		2788		
5		1881		
Ability	38.7**			
high		1590		
low		3078		
Frequency	59.7**		9.6**	
high		2104		1005
low		2564		1064
Orthography	6.1*		11.6**	
redundant		2413		1056
nonredundant		2255		1013
Versatility	84.0**		41.7**	
versatile		2184		996
nonversatile		2485		1073
Decodability	21.0**		26.0**	
easy		2209		1007
hard		2460		1062

(df = 1,20, $p<.05$ = 4.35, $p<.01$ = 8.10) (df = 1,23, $p<.05$ = 4.28, $p<.01$ = 7.88)

quickly matched than their nondistinctive counterparts. For example, the low versatile "ee" in the final position of "tree" makes it distinctive. If no match is found, as with low frequency words, then subunits are examined either to see whether they "look like" word parts or to further processing toward word recognition. In both of these cases versatility would be an asset. On the other hand, children have had less exposure to large numbers of words and are less likely to have developed a set for versatility. They appear to engage in more subunit processing in which orthographic redundancies reflect "wordness" and/or lead to word recognition. Orthographic redundancies are related to the high frequency words to which these children have been most exposed (high redundancy counts often reflect the letter patterns in highly frequent words). It is unclear in the lexical decision task whether subjects normally processed all the way to meaning or frequently made word/non-word judgements on the basis of word "likeness" or word recognition without lexical access.

Picture Task

In the picture task we are sure subjects processed to meaning. In this task, as shown in Table 2, both Grade and Ability again had significant main effects for children with older and/or better readers having faster RT's. For both children and adults there were significant main effects of Frequency, Orthography, Versatility, and Decodability. There were also two significant interactions for both children and adults of Frequency and Versatility (for children, $F(1,20)=4.4$, $p<.05$; for adults, $F(1,23)=14.4$, $p<.01$) and Decodability by Versatility (for children, $F(1,20)=9.1$, $p<.01$; for adults, $F(1,23)=32.8$, $p<.01$) (Figures 3 and 4). Low frequency, nonversatile, and nonversatile words with hard decodable spelling-

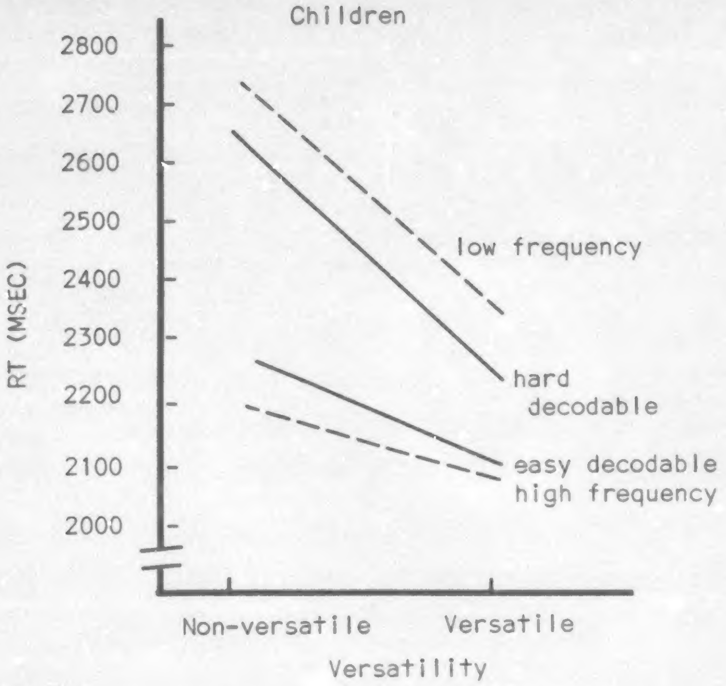


Figure 3. Interactions of Frequency X Versatility and Decodability X Versatility for Children in the Picture Task.

sound patterns have particularly long RT's. Such words have patterns which are harder to match to sound equivalents. Like Johnson (1970), we believe versatility (as it reflects exposure to *different* words with similar patterns) is more influential than orthographic redundancy (as it reflects only number of exposures to certain patterns and can be inflated by a small number of high frequency words) and is more influential in leading to sound generalizations. Whether readers, in particular adult readers, are actually processing to sound or whether the past formation of the generalizations biases readers to quicker processing of versatile words is not clear. However, the impact of Decodability (Table 2; Figures 3 and 4), as it reflects letter-sound correspondences, reinforces the idea that at some level sound equivalents may be involved.

Orthographically nonredundant words were processed quicker than redundant words (Table 2). This result replicates with a new word set, the results of Juel and Solso (1980) for both low and high frequency words. We speculate that words which share certain letter patterns with highly frequent and, therefore, highly redundant words, may be confused with these words. For example, after seeing the highly redundant bigram "th" in initial position in a three letter word the reader expects "the," but if "thy" is encountered, interference, and slower reading times, occur.

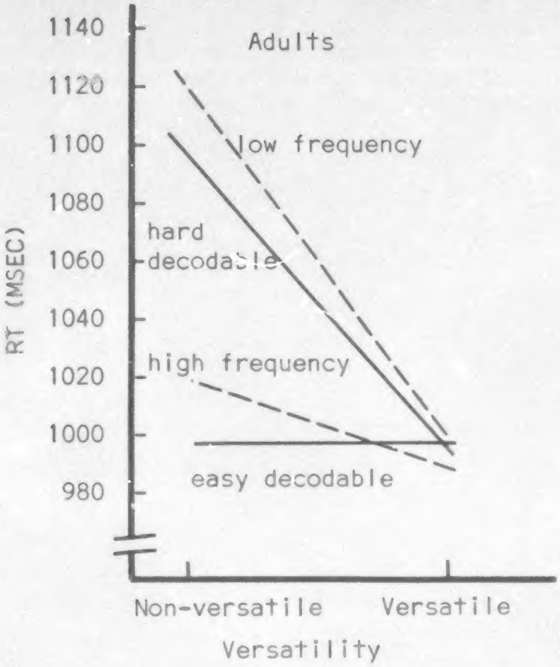


Figure 4: Interactions of Frequency X Versatility and Decodability X Versatility for Adults in the Picture Task.

There were a few significant interactions unique to either adults or children. Adults were found to be especially slowed by redundant, hard decodable words (Orthography X Decodability, $F(1,23)=10.8$, $p<.01$). Such words are less likely to be orthographically distinctive for quick sight word identification, and have less common letter-sound patterns for easy decoding. A Frequency by Decodability interaction for children ($F(1,20)=8.1$, $p<.01$) showed them to be much faster identifying high frequency, easy decodable words than other types—especially low frequency, hard decodable words. This result points to their use of sight word recognition and phonic skills for word identification (See, Juel, 1980). Poor and young child readers were found to be especially slowed by low frequency words (Grade X Frequency, $F(1,20)=4.5$, $p<.05$); Ability X Frequency, $F(1,20)=9.6$, $p<.01$) and nonversatile words (Grade X Versatility, $F(1,20)=7.0$, $p<.05$; Ability X Versatility, $F(1,20)=8.9$, $p<.01$). Both effects likely reflect these children's low sight word vocabularies and subsequent difficulties with words representing unusual letter patterns.

CONCLUSIONS

Our results indicate that Versatility is a highly potent factor in the word identification process. Part of the process of learning to read appears to involve acquiring knowledge about positional letter functions, gained through experience with a number of *different* words. On a practical level our results suggest that the teaching of word identification will be more successful if learners are exposed to a variety of *different* words with similar patterns than exposed to more repeated, but less varied, presentations of a smaller word set, (e.g., In the long run, all else being equal in learning the word "feet," 25 exposures to the word "feet" will be less instructionally effective than 25 combined exposures to "feet," "meet," "feed," etc.).

REFERENCES

Calfee, R. C., Venezky, R. L., & Chapman, R. S. *Pronunciation of synthetic words with predictable and unpredictable letter-sound correspondences* (Tech. Rep. No. 71). Madison: Wisconsin Research and Development Center for Cognitive Learning, 1969.

Carroll, J. B., Davies, P., & Richman, B. *Word Frequency Book.* New York: American Heritage, 1971.

Cattell, J. McK. The time it takes to see and name objects. *Mind,* 1886, *11,* 63-65.

Johnson, D. D. *Factors related to the pronunciation of verbal clusters* (Tech. Rep. No. 149). Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

Juel, C. L. Comparison of word identification strategies with varying context, word type, and reader skill. *Reading Research Quarterly,* 1980, *15* (3), 358-376.

Juel, C. L. & Solso, R. L. Effects of Orthographic and Phonic Structures on Word Identification. In M. L. Kamil and A. J. Moe (Eds.), *Perspectives in Reading Research and Instruction, 29th Yearbook of the National Reading Conference.* Washington, D.C.: The National Reading Conference, 1980.

Mason, M. Reading ability and letter search time: Effects of orthographic structure defined by single-letter positional frequency. *Journal of Experimental Psychology: General,* 1975, *104,* 146-166.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General,* 1979, *108,* 124-197.

Miller, G. A., Bruner, J. S., & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology,* 1954, *50,* 129-39.

Rumelhart, D. E., & Siple, P. Processes of reorganizing tachistoscopically presented words. *Psychological Review,* 1974, *81,* 99-118.

Shankweiler, D. & Liberman, I. Y. Misreading: A Search for Causes. In J. F. Kavanaugh and I. G. Mattingly (Eds.), *Language by Ear and by Eye.* Cambridge, Mass.: The MIT Press, 1972.

Solso, R. L. Positional frequency and versatility of letters for six, seven-, and eight-letter English words. *Behavior Research Methods and Instrumentation,* 1979, *11,* 355-358.

Solso, R. L., Barbuto, P. F. & Juel, C. L. Bigram and trigram frequencies and versatilities in the English language. *Behavior Research Methods and Instrumentation,* 1979, *11,* 475-484.

Solso, R. L., & Juel, C. L. Positional frequency and versatility of bigrams for two- through nine-letter English words. *Behavior Research Methods and Instrumentation,* 1980, *12,* 297-343.

Venezky, R. L. *The structure of English orthography.* The Hague: Mouton, 1970.

MARK AULLS
McGill University

DEVELOPMENTAL CONSIDERATIONS FOR READING RESEARCH: APPLICATIONS TO GOOD AND POOR READER RESEARCH

In cognition and reality, Neisser (1976) argues for a cognitive psychology which is realistic and research which is ecologically valid. Reading represents one form of ordinary human activity which is purposeful and ecologically important to human learning. In the last decade considerable advances have been made in our understanding of reading through the contributions of scholars from various fields of cognitive psychology. However, during this time reading theory and the vast majority of reading research have not given sufficient attention to developmental changes in reading performance or how reading undergoes systematic development. This is particuiarly obvious for research on good and poor readers. Until very recently (Resnick, 1979; Brown, Camprone & Day, 1981; Ryan, 1980) only a very small proportion of good and poor reader research has been conceived within the context of developmental theory.

The purpose of this paper is to argue that there is a serious need to: 1) reanalyze existing assumptions underlying good and poor reader research and 2) reconceptualize approaches to, and the design of future studies of good and poor readers within some form of developmental framework (Doehring & Aulls, 1979; Fischer, 1980; Brown, 1978). The arguments to be presented will begin with a review of the major theoretical assumptions underlying good and poor reader research. This will be followed by a review of the methodology and problems typical of research within each theoretically motivated body of good and poor reader research.

Theoretical Assumptions Underlying Previous Approaches to Good and Poor Reader Research

There have been at least three major approaches to good and poor reader research between 1950 and 1977. Each one has been based on specific assumptions about the nature of the reading process and determinants of reading acquisition.

The first line of research is called the cognitive deficit approach. It is based on the assumption that reading is a learned skill. However skills learning is believed to be controlled by basic cognitive processes which themselves are neurologically determined. Reading deficiency is believed to be caused by a specific deficit in one or more cognitive processes (auditory or visual memory, discrimination, segmentation, blending, etc.)

The second approach is called the subskills approach. It is anchored in the subskills model of reading. This model views reading to be a skill which, unlike language, is not naturally learned through hypothesis-testing. Instead it is viewed to be a complex unitary skill which is consciously learned as a series of subskills. Subskills learning is believed to follow a sequence beginning with simpler subskills and progressing toward the larger, more complex skills.

The third approach is called the difference reader approach. This approach assumes that skills knowledge is not sufficient to explain reading acquisition deficits. The absence of strategies may also lead to deficiencies in reading. This assumption follows from the information processing model of reading. At the time this approach to research was carried out, the most popular information processing model of reading was the analysis by synthesis model (Gibson & Levin, 1975: 1976). In its simplest form a reader makes a hypotheses about the text message (or some segment of it), applies rules to determine what the input would be like if it were true and checks to see if the input is really like that. Hence the reader must not only map knowledge of language onto text but utilize rules and strategies such as prediction, self correction, confirmation, etc. to process text meaning.

The Cognitive Deficit Approach

One of the earliest approaches to the study of poor readers was initiated in the late 1950's. Researchers sought out cognitive deficits to explain severe reading problems. The theoretical assumption underlying this research was that reading deficiency was caused by a specific deficit in one or more cognitive processes. Processes such as visual meaning were thought to control a wide variety of reading and other subskills learning (Kephart, 1960; Wepman, 1967; Johnson & Myklebust, 1967; Kirk & Kirk, 1967; Bateman, 1967; Mann, 1971; Larsen & Hammill, 1975).

A variety of reviews of tests developed to assess cognitive process deficits have been found to be sadly lacking in construct validity and frequently are unreliable for making diagnostic decisions (Sedlack & Weener 1973; Proger et al., 1973; Waugh, 1975; Larsen & Hammill, 1975; Hammill & Wiederhold, 1973; Ysseldyke, 1973; Arter & Jenkins, 1979). More importantly, several thorough reviews of instructional research on various training programs designed to restore cognitive deficits (Minskoff, 1975; Hammill & Larsen, 1974; Silverston & Deichman, 1975; Arter & Jenkins, 1979) have found very meagre evidence of positive effects on reading achievement. One recent comprehensive review of this type of research (Arter & Jenkins, 1979) found that less than 3% of the measures (used in over 100 studies) indicated significant treatment effects. Some authorities suggest that true effects may be considerably lower than obtained effects when poorly designed studies are excluded (Newcomer et. al., 1975).

In summary, three inferences can be drawn from reviews of research, measures and instructional results motivated by the assumption that poor reading stems from cognitive deficits. First, measures of performance cannot be confidently interpreted without accompanying process measures (Ryan, 1981). Second, using intensive design (Moracco & Floyd, 1980) case studies, rather than groups or matched pairs of good and poor readers, may offer a more appropriate basis for studying how highly atypical readers differ from typical readers in skills knowledge, processing strategies and metacognitive functions. Third, there exists little evidence that cognitive deficits preclude growth in reading skills or strategies. Finally, the existing instructional research offers no confidence that cognitive training, distinct from reading skill or strategy training has an influence on reading performance or changes in the poor reader's course of reading acquisition.

The Subskills Approach

A second major approach to research on good and poor readers stems from viewing reading as subskills learning. The major theoretical assumption is that subskills of normal readers are highly interrelated. Hence, the development of one subskill depends upon the development of another but absence of transfer may affect all or many subskills (Gutherie, 1973; Samuels & Schachter, 1978). Some correlational research (Gutherie, 1973) indicates that young and older disabled readers have lower correlations among subskill measures than young good readers who demonstrate highly related subskills. However, survey research attempting to identify developmental subskill learning sequences provides no strong evidence that a hierarchy of subskills exists or that changes in skills knowledge can fully explain changes in reading comprehension or the fluency of oral or silent reading (Athey, 1975). Nonetheless, the likelihood of multiple but typical sequences for subskill learning is still not a completely closed issue (White, 1973; White & Gagne, 1974). Further clozure on this topic awaits the application of new (but available) correlational methods (Doehring et. al., 1981; Fisher, 1980). Lastly, the subskills research methodology has not been coupled with or led to reading training studies where the effects of transfer learning of subskills to reading natural text can be empirically verified.

In conclusion, the direct contribution of learned subskills knowledge to comprehension of text during or following reading has not been shown. In fact, there is no published research to suggest whether readers who exhibit differences in subskill knowledge, such as word recognition, actually perform differently in processing information during text reading (Resnick, 1979). At this time, it can only be concluded that subskills alone cannot fully account for good and poor reader differences (Ryan, 1980) or for differential changes in their course of reading acquisition. Hence studies are needed which utilize multiple strategy and skills knowledge variables in order to describe patterns of reading acquisition within a specified range of ages.

The Difference Reader Approach.

A third approach to the study of good and poor readers has been based on what Cromer (1970) called "the difference-deficit reader distinction." This distinction assumes poor readers use inappropriate strategies for processing text although they may possess a range of skills knowledge similar to those available to skilled readers.

Historically the first line of research focused on good and poor reader strategy differences in organizing words into higher order groups during sentence processing as opposed to assuming a deficit in specific cognitive processes, missing subskills or underdeveloped access to skills knowledge (speed of word recognition). In a series of studies from 1966 to 1971 (Cromer & Wiener, 1966; Oakan, Wiener & Cromer, 1971; Steiner, Wiener & Cromer, 1971) evidence was established that even with similar vocabulary knowledge and accurate word recognition, older poor readers in contrast to good readers could not read words in higher order word groups. Later studies (Isaksen & Miller, 1976, Neville & Pugh, 1977; Rode, 1974; Aulls, 1977; Ryan & Willows, 1979) using oral reading and/or cloze tasks further confirmed that good and poor readers do differ in that poor readers: 1) use or recover syntax cues less often, 2) combine syntactic and semantic cues less often, 3) progress at different rates in acquiring phrase grouping on a broad range of texts and 4) may use different strategies to break through word-by-word reading.

This research offers strong preliminary evidence that deficient subskills knowledge and central cognitive deficits alone cannot account for differences among good and poor comprehenders. It also verifies that strategies can account for differences in reading fluency when word identification skills are controlled. Yet it is still not clear how change occurs in the reader's ability to group words into successively higher order units or to maintain word grouping changes on increasingly difficult texts. Are such changes primarily due to: 1) changes in decoding accuracy and speed (Stenovich, 1980; Samuels & Laberge, 1974) combined with repeated readings of familiar texts (Samuels, 1979) or 2) changes in metalinguistic awareness of and/or strategic use of tex[illegible]tactic cues which signal meaningful word groupings within sentences [illegible] changes in metalinguistic awareness and strategies for processing text through practice in simultaneous listening to and reading sentences in meaningful texts (Chomsky, 1976; Ehri, 1978)?

The second line of research focuses on the observed uses of text cues and strategies during oral reading to construct sentence meaning and between sentence meanings relationships. In the case of cues, the interest is in whether good and poor readers differentially use grapho *H* phoneme cues within words, syntactic cues within words and sentences and semantic cues within and between sentences. The use of strategies refers to self correction strategies, confirmation strategies and the spontaneous self regulation of cues to confirm oral reading mistakes.

Some of the most common findings across studies suggest further insights into text processing differences among good and poor readers (Goodman, 1967; Weber, 1970; Biemiller, 1970; Menosky, 1971; Isaksen & Miller, 1976; Brazee, 1976; Burke, 1976; Leslie, 1980; Pflaum & Bryan, 1980; Fay, Tripin & Townes, 1981). First, poor readers make more mistakes during oral reading which change the sentence and text meaning intended by the author. This appears to be true regardless of text difficulty and control for rate of mistakes per hundred words. Second, poor readers do use context to predict and/or self correct sentence processing mistakes but they are less effective in using context than good readers. The use of self correction using partial sentence context is proportionally similar to or higher than good readers after controlling mistakes per hundred words. However, poor readers seldom use information from the text outside the sentence being processed to self correct. Third, poor readers (beyond the primary grades) use graphic cues as effectively as good readers but they don't coordinate these cues with other cues as well as good readers. Ability to coordinate two or more levels of text cues distinguishes younger good and poor readers. In short, poor readers make more oral reading mistakes, are less aware of the extent to which their mistakes disrupt the meaning they construct during oral reading and are less effective in using multiple cues and alternative strategies in preserving the meaning gist of the entire text.

The difference reader research (and subsequent research focusing on the availability and/or use of strategies) has led to a broader variety of research designs than the other two approaches. Collectively these studies represent the use of cross age samples, control over number of word identification mistakes and text difficulty to allow inferences about differences in good and poor reader reading strategies, comparisons of speech to text processing, and the simultaneous analysis of processing strategies and performance measures. Hence it may be possible to reanalyze this entire body of research through meta-analysis procedures in a manner which will clarify the importance of variables such as higher order word grouping to typical and atypical patterns of comprehension growth displayed among good and poor readers at different age ranges. Also, it may be possible to distinguish the relative contributions of the number and quality of strategies used to the transfer of word grouping across texts which differ in their difficulty. Finally, this body of research has established that strategies play a significant role in text processing and may make a substantial difference in the quality of meaning constructed when reading text.

Methodology and Problems of Good and Poor Readers Research

An initial problem with this body of research is that there is no clear conceptual framework for defining variables such as "good and poor" readers. The selection of good and poor readers typically depends upon standardized norm referenced texts. Usually global reading or IQ scores serve as the basis for defining good and poor readers and to what population the results can be generalized (Valtin, 1979; Golinkoff, 1976). Of itself, this practice makes it quite difficult to compare the results of similar studies which used different tests to define the research sample. Furthermore, the results of some studies may in fact be artifacts of the global ability measures used to define good and poor readers (Reed, 1970) depending on the research design. But most important, the terms good and poor reader may simply be a function of inexperience with specific reading tasks or of age related metacognitive dimensions of growth.

Common Designs Used in Good and Poor Reader Research

Good and poor reader studies typically have used the first three of the following four research designs:

1) Studies which contrast good and poor readers of the same age on specific performance tasks in order to determine differences in skill knowledge. For example, knowledge of words, syntactic cues and main idea units, etc.
2) Studies which compare good and poor readers of the same age on one or more processing strategies. For example, prediction of form class, self correction of mistakes, higher order word grouping of information, inferencing, etc.
3) Studies which correlate knowledge and/or strategies observed among good and poor readers in order to study the differences between group correlation patterns.
4) Studies which provide short training tasks to good and poor readers (using one or more age groups) where training is expected to enhance poor reader performance but not good reader performance given specific hypothesizes regarding differences in access to strategies or skills.

Problems of Interpreting Results.

The first three paradigms for good and poor reader studies offer what Resnick (1979) refers to as a "snapshot" or "steady-state" descriptions of good and poor reader differences or relationships. This type of research is most useful to exclude hypotheses about readers' strategic processing and skill knowledge differences within a given age range or level of reading acquisition. These three types of designs are clearly deficient in characterizing the extent (Fischer, 1980) to which performance differences among good and poor readers can be changed and how important existing differences are to specific reading tasks entailed in successful reading acquisition.

It is not uncommon to find that reading educators interpret the findings of good and poor readers' contrast studies to be generalizable beyond a given age range. One problem with this type of interpretation is that the differences obtained may in fact not be replicable with samples of other poor readers from a different age group. Confident interpretation of the results depends upon subsequent replication of findings across different age group samples or upon contrasting the results to intensive single subject intervention studies (Moracco & Floyd, 1980) replicated with poor readers selected from different age groups.

Suppose reliable and significant differences are found in certain skills or strategy contrasts between good and poor readers of different ages. Must it necessarily be assumed that poor readers are simply less advanced than good readers and both groups will follow the same course of reading acquisition? It is equally possible that the two populations will follow quite different courses of reading acquisition. Furthermore, poor readers within the same experimental group may exhibit very different patterns in the manner that reading skills are related in the number or quality of strategies available for use and in the ability to activate or self regulate available strategies during normal text reading situations. Resnick (1979) and Ryan (1981) suggest that the observed skills or strategies in the poor reader group may simply be symptomatic of generally less adaptive uses for strategies rather than slower but similar development. Finally, it is also possible that the important sources of comprehension failure among poor readers are not due to availability of processing strategies but to lack of awareness that failure has occurred and to inability to self regulate, monitor or check the effectiveness of strategies used (Brown, 1980).

In short, the first three research designs have been widely used in good and poor reader research. Yet significant problems exist in the generalizability of results and how to interpret the results in terms of their instructional and theoretical implications. These are not minor problems. Without serious considerations of whether available research findings hold up under developmental scrutiny, they are not likely to be resolved.

SUMMARY

The collective findings of the three major approaches to the study of good and poor readers offer three central findings. First, it is very unlikely that poor reading performance, regardless of age of the reader, is due to cognitive deficits which control skills knowledge acquisition. Second, skills knowledge deficits may occur but it is the inability to orchestrate this knowledge which is the primary distinction between good readers and both younger and older poor readers. Hence skills knowledge alone is not sufficient to account for changes in text comprehension. Third, absence, disuse or ineffective use of strategies appears to distinguish good and poor readers who have similar reading skills knowledge. Hence the orchestration of skills knowledge during reading in order to become fluent and to recover from comprehension failures may depend heavily on the availability and use of reading strategies.

The major problem with the body of research reviewed is that it lacks a clear developmental framework which suggests how good and poor readers across ages from 6 to 25 might be expected to differ in the availability and spontaneous use of skills knowledge and strategies on similar kinds of materials and reading tasks. One can hardly say what should be learned by poor readers without some sense of how good readers successively change in their reading performance and in the integrated use of knowledge and strategies during text processing and in relationship to multiple criteria for text comprehension. Also it is impossible to infer which differences between good and poor readers may represent long lasting effects on acquisition as opposed to age related short term learning effects. Long lasting differences between good and poor readers may interfere with the pattern of reading development among poor readers while short term differences would not be found among poor readers who are somewhat older.

Third, the absence of a developmental framework for good and poor reader research precludes distinctions between the relative importance of the acquistion of reading skills knowledge and text processing strategies in contrast to metalinguistic knowledge and metacognitive knowledge necessary to many reading comprehension tasks. Recent research reported by Brown (1980); Brown, Campione & Day (1981) and Ryan (1980) clearly suggests that the ability to use metacognitive and metalinguistic knowledge influences success in certain levels of text comprehenion as well as sentence processing and recall. The research of these scholars is based on a developmental framework (Brown, 1978) and entails the use of cross-age research designs, and training studies using intensive designs. In contrast to the research approaches reviewed, developmentally motivated research promises a basis for: 1) excluding alternative hypotheses and 2) identifying what knowledge or strategies poor readers need to learn and the extent to which performance can be changed on specific reading tasks.

In summary, the principle ecological outcomes of good and poor reader studies should be a data base which facilitates sound instructional research hypotheses and ultimately more effective instruction of poor readers. In the classroom sound informal diagnostic observations are among the most critical first steps to sound instruction. Yet improved recommendations for what teachers should look for and how to look simply cannot be validated without adequate information from research about typical and atypical patterns of reading development which reflect the interaction between changes in cognition and changes in specific reading skills knowledge and strategies.

BIBLIOGRAPHY

Anton, J. Intensive experimental designs: a model for the counselor/researcher. *Personnel Guidance Journal,* 1978, *56,* 273-278.

Arter, J. A., & Jenkins, J. R. Differential diagnosis-prescriptive teaching: a critical appraisal. *Review of Educational Research,* 1979, *49* (4), 517-556.

Athey, I. *Essential skills and skills hierarchies in reading comprehension and decoding instruction.* Unpublished paper presented at the conference of National Institute of Education, 1974.

Aulls, M. W. *The acquisition of reading fluency by able and less able readers in grades one, two and three.* Unpublished paper presented at the National Reading Conference, New Orleans, December, 1977.

Biomiller, A. The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly,* 1970,*6,* 75-96.

Bateman, B. Three approaches to diagnosis and educational planning for children with learning disabilities. *Academic Therapy Quarterly,* 1967, *3,* 11-16.

Brazee, P. *A qualitative and quantitative description of eighth grade students' oral reading in both narrative and expository materials.* Unpublished doctoral dissertation, University of Northern Colorado, 1976.

Brown, A. L. Theories of memory and the problems of development: activity, growth, and knowledge.

In L. Cermak & F. I. M. Craik (Eds.), *Levels of processing in memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

Brown, A. L. Metacognitive development and reading. In R. J. Spiro, B. C. Bruce & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension, 1980*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 453-487.

Brown, A. L., Campione, J. C., & Day, J. D. Learning to learn: on training students to learn from texts. *Educational Researcher*, 1981, *10* (2), 14-21.

Burke, E. A developmental study of children's reading strategies. *Educational Review*, 1976, *29* (1), 30-46.

Chomsky, C. When you still can't read in third grade: after decoding what? *Language Arts*, 1976, *53 (3), 288-296*.

Cromer, W. The difference model: a new explanation for some reading difficulties. Journal of Educational Psychology, 1970, *61* (60), 471-483.

Cromer, W., & Wiener, M. Idiosyncratic response patterns among good and poor readers. *Journal of Consulting Psychology*, 1966, *30*, 1-10.

Doehring, D. J., Trites, R. L., Patel, P. G., & Fedorowicz, C. A. M. *Reading disabilities: the interaction of reading language and neuropsychological deficits*. New York: Academic Press, 1981.

Ehri, L. C. Beginning reading from a psycholinguistic perspective: amalgamation of word identities. In F. B. Murray (Ed.), *Recognition of words*. Newark, Delaware: International Reading Association, 1978, 1-32.

Fay, G., Trupin, E., & Townes, B. D. The young disabled reader: acquisition strategies and associated deficits. *Journal of Learning Disabilities*, 1981, *14* (1), 32-35.

Fischer, K. W. A theory of cognitive development: the control and construction of hierarchies of skills. *Psychological Review*, 1980, *87* (6), 477-507.

Gibson, E. J. & Levin, H. *The Psychology of Reading*. Cambridge, Mass: The MIT Press, 1976, 449-453.

Golinkoff, R. M. A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly*, 1975-1976, *11* (4), 623-659.

Goodman, Y. M. *A psycholinguistic description of observed oral phenomena in selected young beginning readers*. Unpublished doctoral dissertation, Wayne State University, 1967.

Guthrie, J. T. Reading comprehension and syntactic responses to good and poor readers. *Journal of Educational Psychology*, 1973, *65* (3), 294-299.

Guthrie, J. G. Models of reading and reading disability. *Journal of Educational Psychology*, 1973, *65*, 9-18.

Hammill, D. D., & Larsen, S. C. The effectiveness of psycholinguistic training. *Exceptional Children*, 1974, *41*, 5-14.

Hammill, D. D., & Wiederhold, J. L. Review of the Frostig visual perception test and the related training program. In L. Mann & D. A. Sabitino (Eds.), *The first review of special education* (Vol. 1). Philadelphia: J. S. E. Press, 1973.

Isaksen, R. L. & Miller, J. W. Sensitivity to syntactic and semantic cues in good and poor comprehenders. *Journal of Educational Psychology*, 1976, *68* (6). 787-792.

Johnson, D., & Myklebust, H. *Learning Disabilities: Educational Principles and Practices*. New York: Grune & Stratton, 1967.

Kephart, N. C. *The slow learner in the classroom*. Columbus, Ohio: Merrill, 1960.

Larsen, S. C., & Hammill, D. D. The relationship of selected visual-perceptual abilities to schooling learning. *Journal of Special Education*, 1975, *9*, 281-291.

Leslie, L. The use of graphic and contextual information by average and below average readers. *Journal of Reading Behavior*, 1980, *12*, 139-149.

Mann, L. Psychometric phrenology and the new faculty psychology: the case against ability assessment and training. *Journal of Special Education*, 1971, *5*, 3-65.

Menosky, D. M. A psycholinguistic analysis of oral reading miscues generated during the reading of varying portions of text by selected readers from grades two, four, six and eight: a descriptive study. Unpublished doctoral dissertation, Wayne State University, 1971.

Minskoff, E. H. Research on psycholinguistic training: critique and guidelines. *Exceptional Children*, 1975, *42*, 136-144.

Moracco, J., & Floyd, M. Intensive designs: research tool for the clinician. *Diagnostique*, 1980, *6* (1), 35-43.

Neisser, U. *Cognition and reality*. San Francisco: W. H. Freeman and Company, 1976.

Neville, M., & Pugh, A. K. Context in reading and listening: variations in approach to cloze tasks. *Reading Research Quarterly*, 1977, *12*, 1-31.

Newcomer, P., Larsen, S. C., & Hammill, D. D. A response to research on psycholinguistic training. *Exceptional Children*, 1975. *42*, 136-144.

Oakan, R., Wiener, M., & Cromer, W. Identification, organization and reading comprehension for good and poor readers. *Journal of Educational Psychology*, 1971, *62*, 71-78.

Pflaum, S. W., & Bryan, T. Oral reading behavior in the learning disabled. *Journal of Educational Research*, 1980, *73* (5), 252-258.

Proger, B. B., Crose, L. H., & Burger, R. M. Construct validation of standardized tests in special education: a framework of reference and application to ITPA research (1967-1971). In L. Mann & D. A. Sabatino (Eds.), *The first review of special education* (Vol. 1). Philadelphia: J. S. E. Press, 1973.

Resnick, L. B. Toward a usable psychology of reading instruction. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1979.

Rode, S. S. Development of phrase and clause boundary reading in children. *Reading Research Quarterly*, 1974, *9*, 603-621.

Ryan, E. B., *Methodological problems in researching strategy deficiencies among good and poor readers: three recommendations*. Unpublished paper presented at the National Reading Conference, San Diego, December, 1980.

Ryan, E. B. Identifying and remediating failures in reading comprehension: toward an instructional approach for poor comprehenders. In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: advances in theory and practice* (Vol. 2). New York: Academic Press, 1981.

Ryan, E. B., Willows, D. M. *The differential utilization of contextual cues by good and poor readers in the middle grades*. Unpublished paper presented to the American Educational Research Association, San Francisco, 1979.

Samuels, S. J. The method of repeated readings. *Reading Teacher*, 1979, *32*, 403-408.

Samuels, S. J., & Schachter, S. W. Controversial issues in beginning reading instruction: meaning versus subskill emphasis. In S. Pflaum-Connor (Ed.), *Aspects of reading education*. Berkeley: McCutchan Publishing Corp. 1978, 44-62.

Sedlack, R. A., & Weener, P. Review of research on the Illinois Test of Psycholinguistic Abilities. In L. Mann & D. A. Sabatino (Eds.), *The First Review of Special Education* (Vol. 1). Philadelphia: J. S. E. Press, 1973.

Silverston, R. A., & Deichmann, J. W. Sense modality research and the acquisition of reading skills. *Review of Educational Research*, 1975, *45*, 149-172.

Steiner, R., Wiener, M., & Cromer, W. Comprehension training and identification for poor and good readers. *Journal of Educational Psychology*, 1971, *62*, 506-513.

Valtin, R. Dyslexia: deficit in reading or deficit in reading research. *Reading Research Quarterly*, 1978-79, *14* (2), 201-221.

Waugh, R. P. ITPA Ballast or bonanza for the school psychologists. *Journal of School Psychology*, 1975, *13*, 201-208.

Weber, R. M. A linguistic analysis of first grade reading errors. *Reading Research Quarterly*. 1970, *5*, 427-451.

Wepman, J. M. The perceptual basis for learning. In E. C. Frierson & B. Barbe (Eds.), *Educating children with learning disabilities: selected readings*. New York: Appleton-Century-Crofts, 1967.

White, R. Research into learning hierarchies. *Review of Educational Research*, 1973, *43*, 361-375.

White, R. & Gagne R. Past and future research on learning hierarchies. *Educational Psychologist*, 1974, *11*, 19-28.

Ysseldyke, J. E. Diagnostic-prescriptive teaching: the search for aptitude-treatment interactions. In L. Mann & D. A. Sabitino (Eds.), *The first review of special education* (Vol. 1). Philadelphia: J. S. E. Press, 1973.

PHILLIP B. GOUGH
University of Texas
at Austin

A COMMENT ON KENNETH GOODMAN[1]

The field of reading owes a considerable debt to Kenneth Goodman. At a time when theoretical discussions of reading seldom went beyond the question of whether words are recognized letter by letter, Goodman (1969/70, 1976) clearly recognized that reading involves much more than word recognition; the reader brings his entire stock of linguistic knowledge to bear on the printed page (cf. Gough, 1975). The now widespread recognition that reading is a psycholinguistic process is largely due to Kenneth Goodman.

At the same time, Goodman began to advocate a particular view of how psycholinguistic knowledge is employed in the reading process. Unhappily, the term PSYCHOLINGUISTIC is sometimes equated with that view. All of us should recognize that the fact that reading is psycholinguistic does not entail that reading is a guessing game. For proof, we need only turn to the model of skilled reading offered by Gough (1972). No less than half of its hypothetical stages and processes directly involve the reader's linguistic knowledge; it could well be called a psycholinguistic model. Yet this model attributes to the skilled reader no expectations, hypotheses, or predictions. (In fact, Gough, Alford, and Holley-Wilcox, 1979, have argued that "psycholinguistic guessing" would interfere with, rather than promote, fluent reading.) Reading IS psycholinguistic; whether it is "top-down," or "bottom-up," or "interactive" is a separate question.

But I believe that Goodman also deserves our praise for vigorously defending an extreme position on this question. It is currently popular to adopt an interactive view of reading, as if this were a kind of wise synthesis of the foolish extremes that reading is bottom-up or top-down. I do not see it this way. To my mind, defending the interactive view must be (to borrow a line from Darrell Royal) like kissing your sister; it is certainly not wrong, but neither can it be very interesting.

There is no question that reading is interactive at some level: the blind cannot read, and the sighted reader who does not know Polish cannot read that language. The reader's linguistic knowledge MUST combine with information arising from the printed page to determine meaning, so there must be interaction somewhere in the reading process. But I subscribe to the view that scientific progress is made through the disconfirmation of precisely false hypotheses, not the demonstration of vague truths (cf. Platt, 1964).

As I see it, the only interesting question about the interaction of visual and linguistic information is how far up the visual goes before it is affected by the linguistic, or (equivalently) how far down the linguistic can reach to affect the visual. So while I will argue that Goodman is wrong, I would praise him for defending a provocative, and thus heuristic, view on this matter.

I would finally praise Goodman for his empirical investigations (e.g., Goodman and Gollasch, 1980). His insightful observations of oral reading errors (what he terms MISCUES) committed by children ranging from Appalachia to Hawaii have considerably enriched our data base and increased our appreciation of linguistic variation in reading acquisition. Moreover, his crucial observation that many (if not most) children's miscues cannot be attributed to graphemic or visual confusions clearly demonstrates that a model like Gough (1972) does not accurately describe the beginning reader. The child who reads aloud "Little Miss Muffet sat on a chair" is plainly doing something other than bottom-up processing.

---

[1] This paper contains the substance of my opening remarks in a debate with Kenneth Goodman at the National Reading Conference, San Diego, December 4, 1980.

A model like mine must be revised if it is to explain such observations. My current inclination is to revise it like this. Suppose we admit that those psycholinguistic processes in the reader which construct syntactic and semantic representations (processes which I subsumed under the eponym MERLIN, in order to suggest the wonder of their feasts) can, in the absence of any incoming information, generate an hypothesis about what comes next. (It is certainly the case that when a speaker pauses or stops in the middle of a . . . . . . sentence, we often do anticipate a particular word.) Let us further suppose that the speed with which this hypothesis is generated depends upon the extent of syntactic and semantic constraint provided by the prior context, in short, with the PREDICTIVENESS of the context to that point. (It is important to note that the predictiveness of context is related to, but not identical with, the PREDICTABILITY of a target word. A highly predictable word must be preceded by predictive context, but a predictive context need not be followed by a predictable word: "Now is the time for all good men to come to the aid of their children.")

Now suppose we imagine that this process is completely independent of the reader's current bottom-up processing: when the reader's eye reaches a given point in text, both Merlin and the bottom-up process begin—independently—to select a word; whichever yields a word first will determine what word is read. It should be obvious that this model, unlike its predecessor, WOULD predict the occurrence of miscues which are semantically and syntactically, but not graphemically, consistent with the prior context. In fact, it predicts that such miscues will occur whenever Merlin predicts a word other than the target word before that word can be decoded. This should never happen if decoding is fast, but if decoding is slow (for any reason), then Merlin may come up with a word before the bottom-up process does. If Merlin's word is not the word on the page, a semantically and syntactically appropriate miscue will be the result. The probability that this will happen will depend directly on the predictiveness of the context, and with the decoding difficulty presented by the target word.

This revised model, unlike its predecessor, might account for the miscues of Goodman's subjects, and it would do so without any interaction whatsoever: it assumes that top-down and bottom-up processing proceed in complete independence. It remains to be seen whether such a model is viable. It is certainly falsifiable, for it clearly must predict that the incidence of oral reading errors will vary directly with the predictiveness of text preceding unpredicted target words, and with anything (e.g., word length) which retards decoding. Moreover, it entails that the time it takes to read a word correctly will decrease with its predictability. But whether or not this model is correct, we must concede Goodman's basic point: the child's oral reading is clearly the product of both top-down and bottom-up processes, either independently (as I here suggest) or interactively (as Goodman surely would have it).

Whichever is correct, it would seem that Goodman and I could agree that the child's reading is the result of some combination of top-down processing, bottom-up processing, or the interaction of the two. It would be hard to deny that the ability of the skilled adult might be partitioned in the same way. But then it follows that the DIFFERENCE between the beginning and the skilled reader must be found in one or more of the same three factors.

That difference is considerable. To illustrate this point, we asked 12 second graders (drawn from a Cub Scout den) and 12 freshmen (drawn from the University of Texas at Austin) to read aloud a 345-word children's story entitled "Grandma and the Skunks." On the average, the second graders took 314 seconds (an oral reading rate of 66 words per minute) and committed 37.83 miscues. The feshmen read the same stroy in a mean of 97 seconds (a rate of 212 words per minute), but committed only 3.42 errors. (It is perhaps worth noting that the college students' error rate on "Grandma and the Skunks" is not conspicuously different from the error rates reported by Stevens and Rumelhart, 1975, and Alford, 1980, for college students reading college-level texts.)

The point is that the oral reading of skilled adults is rapid and (nearly) error-free; by comparison, that of the second grader is slow and faulty. In order to attain the college students' level of competence, the second grader must triple his speed, while at the same time reducing his errors ten-fold. The question is, how can he do it? Where can he improve?

We must find the answer among three mutually exclusive and exhaustive candidates: the top-down, the bottom-up, and their interaction. Let us consider them in turn.

It may well be that there is a difference between the novice and the skilled reader in top-down processing, in the ability to draw upon prior context to anticipate a forthcoming word. But there is an upper bound to improvement here, and it is severely limiting: even the most skilled adult reader can correctly anticipate no more than one word in four (Gough, Alford, and Holley-Wilcox, 1979). Moreover the words he CAN anticipate are those he needn't, for they tend to be short, frequent, easy-to-recognize words (mostly function words), while the words which cannot be anticipated are just those words which present the greatest decoding problems (Alford, 1980). Top-down processing cannot possibly account for the rapid, error-free reading of the skilled adult. So we must look for improvement in either bottom-up or in interactive processing to make the beginner skilled.

Happily, there appears to be a way to arrange a direct confrontation of these alternatives. Suppose we were to compare more skilled with less skilled readers (either older with younger, or good with poor readers of the same age) in their ability to recognize isolated words. We will surely find a difference; most achievement tests make direct use of this difference to MEASURE reading ability. Clearly, this difference must reflect a difference in decoding ability and only that ability, for the recognition of isolated words can scarcely depend on anything else.

Now suppose we compare the same groups of readers in recognizing exactly the same words presented in context. Here it seems clear that both bottom-up and interactive processing (and even top-down) can contribute to performance. Because of the established difference in bottom-up processing, we surely must expect at least as large a difference in recognition in context as we found in isolation. But if there is also a difference between good and poor readers in interactive processing, or if the difference in interactive processing is more important than the difference in decoding, then the difference between them in recognizing words in context must be even larger than that obtained with isolated words.

This prediction is blatantly false. Stanovich (1980) has reviewed nearly a dozen studies which have compared good and poor readers in their recognition of words in isolation and in context. In every instance, there is a significant difference between the groups in recognizing isolated words. But in no instance has the addition of context been found to increase that difference. Indeed, in many studies, context has been found to reduce the difference between good and poor readers. This is not to say that context does not facilitate word recognition: it does, for both groups. But if the greater speed and accuracy of the skilled reader are to be ascribed mainly to a superiority in interactive processing, then context should help the skilled reader more than the unskilled. This is not the case: context consistently benefits the poor reader more than the good.

I conclude that Goodman is dead wrong about what separates the skilled adult from the beginning reader, and hence about what must be accomplished in reading acquisition. The most conspicuous difference between good and poor readers is found in the swift and accurate recognition of individual words, in decoding, and the mastery of this skill is at the heart of reading acquisition (cf. Gough and Hillinger, 1980). But it should not be inferred from this that I completely disagree with Goodman's views on reading instruction. I believe that Goodman's insistence on reading for meaning is exactly right. Our problem is to find a way to teach the child to decode while doing just that.

REFERENCES

Alford, J. A., Jr. *Lexical and contextual effects on reading time.* Unpublished doctoral dissertation, University of Texas at Austin, 1980.

Goodman, K. S. Reading: a psycholinguistic guessing game. In H. Singer and R. R. Ruddell (Eds.), *Theoretical models and processes in reading* (2nd Ed.). Newark, DE: I.R.A., 1976. pp. 497-508.

Goodman, K. S. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly,* 1969/70, *5,* 9–30.

Goodman, K. S., and Gollasch, F. V. Word omissions: deliberate and non-deliberate. *Reading Research Quarterly,* 1980, *16,* 6–31.

Gough, P. B. One second of reading. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Mass: M.I.T., 1972.

Gough, P. B. The structure of the language. In D. D. Duane and M. B. Rawson (Eds.), *Reading, perception, and language.* Baltimore: York Press, 1975.

Gough, P. B., Alford, J. A., Jr., and Holley-Wilcox, P. Words and contexts. In M. L. Kamil and A. J. Moe (Eds.), *Reading research: studies and applications.* Clemson, South Carolina: National Reading Conference, 1979.

Gough, P. B., and Hillinger, M. L. Learning to read: an unnatural act. *Bulletin of the Orton Society,* 1980, *30,* 179–196.

Platt, J. R. Strong inference. *Science,* 1964, *146,* 347–353.

Stanovich, K. E. Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly,* 1980, *16,* 32–71.

Stevens, A. L., and Rumelhart, D. E. Errors in reading: analysis using an augmented transition network model of grammar. In D. A. Norman and D. E. Rumelhart (Eds.), *Explorations in cognition.* San Francisco: Freeman, 1975.

DONALD J. LEU, JR.
Syracuse University

QUESTIONS FROM A METATHEORETICAL PERSPECTIVE: THE INTERDEPENDENCE OF SOLUTIONS TO ISSUES INVOLVED IN THE DEVELOPMENT OF READING COMPREHENSION MODELS

INTRODUCTION

In the field of reading, a major development during the last decade has been the generation of numerous theoretical models describing the comprehension process. Typically, investigators have constructed models of the reading process based on a set of specific, empirical results, often from their particular paradigm, methodology, or investigative approach (cf. Goodman, 1976; Gough, 1972; LaBerge & Samuels, 1974). It is only recently (Beaugrande, 1981; Danks, 1978) that attention has been directed at the general, underlying characteristics of comprehension models. This paper will argue that it is critically important to the future of comprehension models that we begin to view them from a larger, metatheoretical perspective. We need to use the available models themselves as a data set in order to investigate two essential questions: 1) What are the major issues that models of reading comprehension are currently attempting to resolve? and 2) What are the possible solutions to each issue? This paper will present a set of answers to both questions, organized in the form of a taxonomy.

Specifically, this paper will propose that an adequate model of reading must address at least six fundamental issues:

1. The temporal organization of central processing components.
2. The nature of the component processing relationships.
3. The direction of processing.
4. The amount and type of control that exists over processing.
5. The method of lexical access.
6. The product(s) of the system.

Within each issue, it will define a set of possible solutions.

This larger, metatheoretical perspective reveals a particular insight crucial to the development of reading comprehension models: the solutions to four major issues are largely dependent on whether or not central processing components are temporally ordered. Because the issue of temporal organization seems to limit the possible solutions to other issues, the paper will argue that we need to focus greater attention during the next decade on understanding the temporal organization of central processing components.

ISSUES AND SOLUTIONS: AN INITIAL TAXONOMY FOR READING MODELS

The Temporal Organization of Central Processing Components

Every model must somehow address the issue of temporal organization. It must answer two questions: 1) Are the various components temporally ordered? and 2) If so, what should the nature of that ordering be?

If the components that exist in central processing are temporally ordered, some sort of *serial* processing is necessary. If they are not, processing may take place in a *parallel* manner. With serial processing one component transforms the available data and passes it on to a second component. This second component uses the transformed information and in turn transforms it, passing it on to the next, temporally-ordered component. Models such as those proposed by Gough (1972) and Massaro (1975) are typical examples.

Depending upon the nature of the temporal ordering, several types of serial processing are possible: pure serial, serial by-pass, and serial return. In a *pure serial* system (cf. Gough, 1972) three conditions apply: 1) Central processing occurs in a consistent number of stages without any by-pass options, 2) There is a strict, unidirectional flow to the processing of information without any return loops, and 3) All data is directly transformed by each processing component once and only once.

A *serial by-pass* system differs from pure serial system in only one respect: information is not processed in a consistent number of stages. The system is somewhat flexible since information has the opportunity to by-pass one or several processing components (cf. LaBerge & Samuels, 1974).

Flexibility in central processing can also be achieved with a *serial return* system. Here processing is returned to an earlier level if the results fail to meet certain conditions at a later level. A decision rule is used to determine whether or not these conditions are met. Goodman (1976a; 1976b) proposes a serial return solution to the issue of temporal organization in order to capture hypothesis-testing behavior.

The various components at the cognitive and linguistic center of a model of reading processing do not have to be temporarily ordered. The issue of temporal organization is solved in a parallel processing model by having the various components operate simultaneously on the same unit of input. A parallel model (cf. Rumelhart, 1976) has the separate component processes typical of a serial model but it lacks a series of temporally ordered stages. Everything supposedly happens at once to the fixed unit of input that enters central system processing.

The Nature of Component Processing Relationships

A second issue that a model of reading must resolve concerns the type(s) of component processing relationships that occur. How, in other words, does the processing in one component effect the processing in others? Four solutions to this issue are possible: there may be feedback, autonomous, interactive, or independent component processing relationships.

Serial models may have feedback or autonomous relationships between processing components. In a *feedback* system the results from a later level are used continuously, without decision rules, to assist the processing at other, earlier levels (cf. Massaro, 1975).

If a serial return system does not have feedback loops, relationships between components are *autonomous*. In an autonomous system, later components depend on the results from processing in earlier components for their input data (cf. Gough, 1972). Each component does not operate on the same raw data but, instead, is dependent on the output from the immediately preceeding component.

Either of the component relationships (feedback or autonomous) may combine with any of the serial processing types (pure serial, serial by-pass, or serial return). It is also possible for both feedback and autonomous relationships to combine in the same serial model. This is due to the fact that the direction of feedback and autonomous relationships do not co-occur.

Parallel models may have either interactive or independent relationships between processing components. *Interactive* relationships in a parallel system perform a function somewhat similar to feedback relationships in a serial system; the results from any one processing component are used to facilitate the processing in other components. There is an important difference, though, between feedback and interactive relationships. Since feedback relationships assume temporal ordering of component processes, the results from any single component may only assist the processing in components that precede it in temporal order. Interactive relationships do not assume temporal ordering of component processes. The results from any component in an interactive system may therefore assist the processing in *all* other components (cf. Rumelhart, 1976).

If parallel systems are not interactive they are *independent.* When independent relationships exist, no component influences any other component during the simultaneous processing taking place in each. The final result from each component must somehow be integrated, but this occurs after separate, independent processing. Note that this is not the same as the autonomous component processing relationship described earlier. Autonomous components depend on the output from earlier components and are not, therefore, completely independent. The logogen model suggested by Morton (1970) seems to have largely independent component processing relationships, especially between the visual system and the logogens.

It is essential to understand that the type of component processing relationships that may exist depends on a model's solution to the issue of temporal organization. One that is temporally ordered and uses some form of serial processing (pure serial, serial by-pass, serial return) is limited to having feedback and/or autonomous relationships between processing components; it may not have interactive or independent relationships. Conversely, a parallel system may have interactive or independent relationships between processing components; it may not have feedback or autonomous relationships. The reason for this dependency should be clear. Feedback and autonomous relationships assume temporal ordering of system components while interactive and independent relationships do not.

The Direction of Processing

An adequate model of reading must also resolve a third issue—the direction of processing. Two questions must be answered: 1) Does directionality exist during processing? and 2) If so, does the temporal ordering of component processes proceed from surface features to abstract representations or from abstract representations to surface features? Once again, the solution to this issue depends on how one resolves the question of temporal organization. A serial system must have directionality during component processing since the components, it is assumed, are temporally ordered. A parallel system, at least during central processing, may not have directionality since component processes are not temporally ordered.

Directional systems may be either top-down or bottom-up. A *top-down* system temporally orders components from those that process abstract representations to those that process surface features. There are no completely top-down models. Models by Goodman (1976a; 1976b) and Smith (1971) are, however, most frequently associated with this approach. In a *bottom-up* system, components are ordered from those that process surface features to those that process abstract representations (cf. LaBerge & Samuels, 1974).

Control of System Processing

The fourth issue that an adequate model of reading comprehension must resolve has to do with the control of system processing. Is there some sort of "executive" (Pascual-Leone & Goodman, 1977) or "homunculus" (Morton & Broadbent, 1967) that takes charge and directs the flow and amount of mental energy during processing? If there is, the system is controlled, if not it is uncontrolled. Unfortunately, few current models propose solutions to this issue.

The defining characteristic of a *controlled* system is its flexibility. Cognitive effort is directed by a "supercomponent" to meet different task demands (skimming, searching, pleasure reading, critical reading, editing, etc.) and to adjust to the difficulty level of any particular task.

There are at least three possible types of controlled systems: controlled/reorder, controlled/director, and controlled/mixed. In a *controlled/reorder* system, a reader achieves flexibility by having the ability to reorder the temporal sequence of processing components. Depending upon the task, processing could be largely top-down or bottom-up. In a *controlled/director* system, a reader achieves flexibility by directing the amount of mental energy going to higher or lower levels of processing. *Controlled/mixed* systems combine both controlled/reorder and controlled/director functions.

The particular type of control system possible in any model depends again on the solution to the issue of temporal organization. A controlled, serial model may use any of the three types of controlled systems. A controlled parallel model may only use a controlled/director system.

Serial models may use a controlled/reorder system since they assume temporal ordering of component processes, i.e. if a serial model contains an ordered series of processes it may also reorder them. A serial model may also use a controlled/director system since the only assumption of this type of control is that separate component processes exist, an assumption met by a serial model. Because serial models may have either controlled/reorder or controlled/director systems, they may also have a controlled/mixed system.

Parallel models may only have a controlled/director system. They may not have a controlled/reorder system since parallel processing assumes that there is no temporal ordering of component processes. Because controlled/reorder systems may not occur in parallel models, controlled/mixed systems may also not occur.

An *uncontrolled* system lacks a "supercomponent" to direct processing. There is no way to alter existing pathways just as there is no way to direct greater or lesser cognitive effort to any particular processing component. Either serial or parallel systems may be uncontrolled. Most current models appear to lack control of central processing.

Lexical Access

The fifth issue addressed by an adequate model of reading comprehension is the method of lexical access. There are three possible means of lexical access: indirect, direct, and mixed.

When phonological recoding must take place between the graphic information on the page and a word's lexical meaning in the head of a reader, indirect lexical access exists (cf. Gough, 1972).

When no intermediate step exists between the processing of graphic sources of information and lexical meaning, we have *direct lexical access;* phonological recoding is not necessary for any item in the text (cf. Smith, 1971).

A *mixed lexical access* system is one that can be both direct and indirect, lexical access can take place with or without phonological recoding (cf. LaBerge & Samuels, 1974).

It should be pointed out that each of these methods of relating visual information and lexical meaning are not necessarily unidirectional, going only from visual information to lexical meaning. Rather, it is also possible for one to go from lexical meaning to a visual representation in order to check out a hypothesis initiated by higher order processing. The nature of the interaction between text and lexical knowledge, therefore, is independent of solutions to the issues of temporal organization or directionality.

The System Product(s)

The solution to the final issue is also independent of other issues. It concerns the nature of the product(s) of the system. What, in the final analysis, does the processing model accomplish? Three outcomes are possible: comprehension, production, or both comprehension and production. This particular issue, on the surface, does not seem to be very important. Reading can, of course, result in both comprehension and production. One would expect, therefore, that models of reading would naturally have both outcomes as possible system products. Unfortunately, most incorporate only comprehension in a model that depicts silent reading. Seldom do models attempt to demonstrate the relationship between production and comprehension that occurs during the oral reading process.

CONCLUSIONS

It is clear from this taxonomy that the possible solutions to a number of critical issues depend on the prior solution to the issue of temporal organization. Figure 1 summarizes the issues and the possible solutions for systems with and without temporally ordered central processing components. Of the six issues important to the construction of an adequate model of reading processing, four are completely or largely dependent on whether or not there is temporal ordering among the various processing components: temporal organization, component processing relationships, directionality, and control. Only solutions to issues of lexical access and system product(s) are not dependent on the solution to the issue of temporal organization.

When the dependency relationships between these theoretical issues are understood it is clear that the question of temporal organization should receive a high priority. Its resolution will define the answers to a number of other questions. In the past decade, however, many of the resources and much of the discussion in the area of reading comprehension has focused on the issue of directionality. The attention paid to the issue of directionality has caused the issue of temporal organization to go virtually unnoticed.

FIGURE 1 A Taxonomy of Issues Critical to the Formulation of a Model of Reading Processing Organized by the Solution to the Issue of Temporal Organization

Issue	Solutions Possible With Temporally Ordered Processing Components	Solutions Possible With Non-Temporally Ordered Processing Components
Temporal Organization	Serial —Pure Serial —Serial By-Pass —Serial Return	Parallel
Component Processing Relationships	Feedback Autonomous	Interactive Independent
Control	Controlled —Controlled/Director —Controlled/Reorder —Controlled/Mixed Uncontrolled	Controlled —Controlled/Director Uncontrolled
Directionality	Directional —Top-down —Bottom-up	Nondirectional
Lexical Access	Direct Access Indirect Access Mixed Access	Direct Access Indirect Access Mixed Access
Product(s)	Comprehension Only Production Only Both Comprehension and Production	Comprehension Only Production Only Both Comprehension and Production

Instead of continuing to commit much of our effort to an issue with an extensive research history and relatively limited utility (in terms of helping to solve other issues) we should seek out those issues that, in their resolution, limit the range of possible solutions to other issues. The financial and political realities of the current decade mandate that the extremely limited resources available for basic research be intelligently directed. We can little afford the luxury of investigating issues with only a single theoretical pay off. This suggests that, at the present time, the issue of temporal organization possesses the greatest utility for researchers. Solutions to a number of other issues are limited by its resolution.

Unfortunately, little work has been done in the area of temporal organization. Some initial attempts are underway (Fears & Danks, 1977; Just & Carpenter, 1980) but the work is far from providing conclusive evidence for a resolution to this important issue. Hopefully, research will direct greater attention to this critical question during the 1980's.

ACKNOWLEDGEMENT

The author would like to thank Herbert D. Simons, Robert B. Ruddell, Charles Kinzer, and Karen Carroll for their comments after reading an earlier version of this paper.

REFERENCES

Beaugrande, R. Design criteria for process models of reading. *Reading Research Quarterly*, 1981, *16*, 261-313.

Danks, J. H. Models of language comprehension. *Polish Psychological Bulletin*, 1978, *9*, 183-192.

Fears, R. & Danks, J. H. *Levels of processing in oral reading*. Paper presented at the meetings of the Midwestern Psychological Association; Chicago, May, 1977.

Goodman, K. S. Behind the eye: what happens in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading*. Newark, Delaware: IRA, 1976a.

Goodman, K. S. Reading: a psycholinguistic guessing game. In H. Singer & R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading*. Newark, Delaware: IRA, 1976b.

Goodman, K. S. and Goodman, Y. Learning to read is natural. In L. Resnick and P. Weaver (Eds.) *Theory and Practice of Early Reading*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1980.

Gough, P. B. One second of reading. In J. F. Kavanaugh and I. G. Mattingly (Eds.), *Language by Ear and by Eye*. Cambridge, Mass.: MIT Press, 1972, 331-358.

Gough, P. B. and Cosky, M. J. One second of reading again. In N. J. Castellan, Jr., D. P. Pisoni, and G. R. Potts (Eds.), *Cognitive Theory Volume 2*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Just, M. A. & Carpenter, P. A. A theory of reading: from eye fixations to comprehension. *Psychological Review*, 1980, *87*, 329-354.

LaBerge D. & Samuels, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 1974, 293-323.

Massaro, D. W. Primary and secondary recognition in reading. In D. W. Massaro (Ed.), *Understanding Language*. New York: Academic Press, 1975, 241-290.

Morton, J. A functional model for memory. In D. A. Norman (Ed.) *Models of Human Memory*. New York: Academic Press 1970.

Morton, J. & Broadbent, D. E. Passive versus active recognition models, or is your homunculus really necessary? In W. Wathen-Dunn (Ed.) *Models for the Perception of Speech and Visual Form*. Cambridge, Mass.: MIT Press, 1967.

Pascual-Leone, J. & Goodman, D. *Intelligence and experience: a neo-Piagetian approach*. Unpublished manuscript, York University, 1977.

Rumelhart, D. E. *Toward an interactive model of reading*. Center for Human information Processing, University of California, San Diego. La Jolla, California. Report Number 56, March 1976, 40 pp.

Smith, F. *Understanding Reading*. New York: Holt, Reinhart & Winston, 1971.

EDWARD FRY
Rutgers University

A PARTIAL READING MODEL UTILIZING LANGUAGE UNIT SIZE BY FREQUENCY

The unit size problem has been around for some time. Mathews (1966) reports that ancient Greek and Roman children were taught to read by first mastering letters, then syllables, then words and finally sentences. But just because various units have been around for some while doesn't mean that we know which unit to lean upon. Chall in 1967 found a "Great Debate" still going on between phonics, presumably graphemic units, and whole words. Gleitman and Rosen (1973) have rekindled some interest in the syllable as a teaching unit. The syllable has fallen into some disrepute from the once high honor it held when America was an infant nation—both the New England primer and the Webster Spellers emphasized the syllable as a major beginning literacy unit. Even the famous McGuffey readers, used so heavily in the last century, introduced new words in a syllabicated form, though McGuffey also advocated both phonics and the whole word method.

Modern reading series tend to have a goodly dose of phonics laid on top of a whole word beginning reading approach, plus a pious reverence for comprehension at all levels. In the middle of this century the dose of phonics was considerably lighter. The trends on emphasis or de-emphasis of various unit sizes are at best not well substantiated by research results.

Likewise, modern reading theories which are often concerned more with the "reading process" of either mature or immature readers show no clearcut concern with unit size. Often the word is seen as the fundamental unit being processed, or else the fundamental process is really very abstract such as a schema or past experience. Hence the graphic unit is not of great importance. However, some theories do concern themselves with individual letters or even parts of letters, then they infer a growth of unit sizes. LaBerge and Samuels (1974) postulate that smaller units achieve "automaticity" when larger units are used. Venezsky and Calfee (1970) postulate the interest concept of LMU or Largest Manageable Unit being used by the reader.

This paper will not attempt to take sides on the "optimal" sized unit but rather attempt by simple graphic progression to layout the various sizes of graphical units. But its major purpose is not to state the obvious, namely that there are graphical units of increasing size and complexity, but that most of these units are amenable to a basic research technique called frequency counts. This is not new information but it is basic information that is often ignored or slighted by both reading curriculum developers and theoretical reading process model builders.

The history of word counts is shrouded in antiquity. It is possible that in 1588 Timothe Bright utilized one in developing a shorthand system and probably one was used in 1721 in developing an English dictionary (DeRocher et al 1973). In more modern times, Ayres (1915) counted 400,000 running words as a background for his *Measuring Scale for Ability in Spelling*. But all our modern research in word counting is based more upon the work of Thorndike in 1921 which was published in his classic *Teacher's Word Book of 10,000 Words*. The research in words has continued up to relatively recent times with Kucera and Francis' *Analysis of Present Day American English* in 1967 and John Carroll's *American Heritage Word Frequency Book* in 1971.

Frequency counts of units smaller than words have also been important. Cryptographers from Edgar Allen Poe to modern times have utilized letter frequency counts, though unfortunately the designers of the present typewriter keyboard did not take them into account before the present keyboard became so standardized, nor apparently do authors of penmanship books.

Modern phoneme-grapheme counts stem largely from the work of Dr. Paul Hanna at Stanford University who, in 1951, guided a student named Thomas Moore in a hand phoneme-grapheme frequency count of 3,000 words. A much larger study by Hanna, Hanna, Hodges, and Rudorf was done with USOE sponsorship in 1967. This study is partly summarized in the book *Spelling, Structure and Strategies* (Hanna, Hodges, and Hanna 1971). The Hanna and Hodges frequency count of phoneme-grapheme correspondence is comparable in size and importance to the Carroll word count.

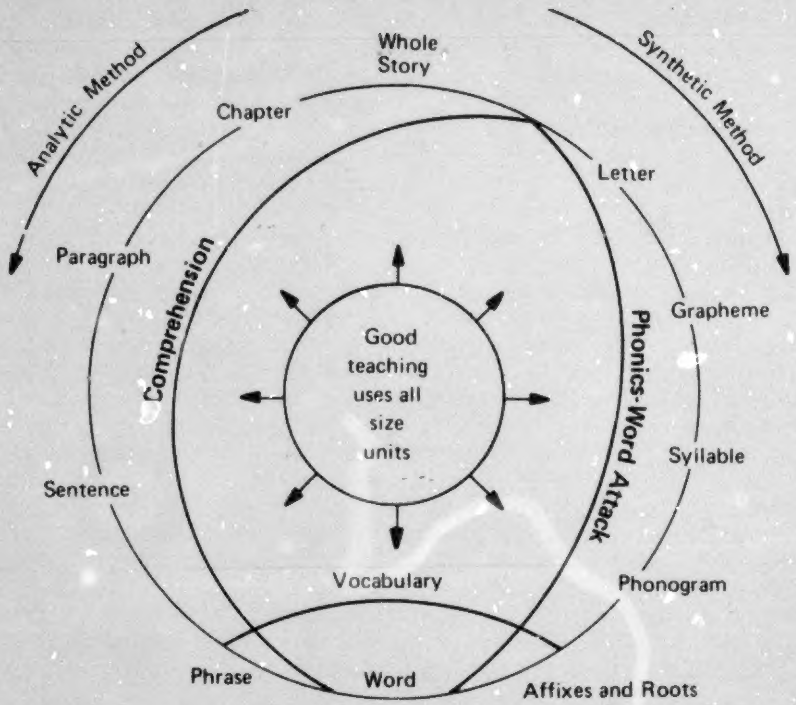
A third major frequency count is the syllable count done by Sakiey (1977). Earlier work in syllable counting is at best fragmentary and at worst unobtainable. The best early work was done by Dewey (1923 and 1950). He counted not graphic syllables but phonetic or speech syllables. However some work in counting grahic syllables was done by Dolch (1940) and Osburn (1954) but they were pre-computer studies. Sakiey not only had computers available but a massive background of data—the Carroll count of 5 million running words with which to "weight" her syllable count. Hence there is excellent information available to answer the simple question "Which are the most commonly read syllables?"

There are very few frequency counts of reading units larger than a word. Most of the work in this area seems to be in identifying the unit sizes and in determining their importance in the reading process. For example, the work now going on in figurative language, story grammars, schema, and discourse analysis is concerned both with graphic units and with cognitive units. Increasingly there is evidence not only that larger-than-word units exist, but that they can be graded; that is, that some are easier than others. Frequency counts that objectively determined "commonness" would be highly desirable.

Hopkins and Moe (1978) did have limited success in attempting to count word strings or groups of 2, 3, 4 words utilizing a computer and text from childrens' trade books used in primary grades. They found some commonness for 2 word strings, a little for 3 word strings and almost none for 4 word strings. This research shows that counting of some units larger than words can be done but it is difficult both from data collection and utility conceptualization.

Perhaps for the sake of completeness we should illustrate the various graphic units that are encountered by the reader. They are: Letters, Graphemes, Clusters (such as blends), Syllables, Phonograms, Affixes and Roots, Words, Phrases, Sentences, Paragraphs, Chapters, and Whole Stories. This list could possibly be extended a bit, for example at the lower end below the "letter" we might have some "distinctive features" and at the upper end beyond the "whole story" we might have the trilogy or series of books, but the basic list will serve as a basic core of graphic unit sizes.

The partial reading model is proposed here simply because there are different-sized language units, and all of these different-sized units have various frequencies of occurrences. For example, there are more and less common graphemes, more and less common syllables, words, etc. *The familiarity or commonness of the unit has an effect on the reading process of the mature*


Analytic Method
Whole Story
Synthetic Method
Chapter
Letter
Paragraph
Comprehension
Grapheme
Good teaching uses all size units
Phonics-Word Attack
Syllable
Sentence
Vocabulary
Phonogram
Phrase
Word
Affixes and Roots

reader and on the learning to read process. Most other "reading models" have failed to take this factor into account.

One important measure of commonness is frequency of occurrence in large samples of written material. This is admittedly a group measure and more suitable for groups or "typical" readers. A more ideal but nearly impossible task would be to determine commonness for a particular individual.

Authors of educational curriculum materials have often used frequency in constructing materials. Some basal reading series and many remedial materials are based on high-frequency word counts; phonics materials and spelling programs have also been influenced by frequency studies. Yet reading theoreticians have often ignored frequency and left it out of their "models." An unfortunate consequence of this is that language frequency studies are not "theory based" and tend to fall from favor by both researchers and scholarly journal editors.

Now before some "top down" theorist such as Goodman, or even some interactivist like Rummelhart, attacks the graphic unit as hopelessly "bottom up", let me waive a flag of truce and hide behind the title which states that "unit size by frequency" is only a *partial* theory. Its claim to academic respectability is that it is objective and verifiable, something some other theories are not.

If one wishes to look at either the reading process or learning to read as more, or mostly, cognitive I certainly have no objections. In fact if someone wishes to do a frequency count of schema or story grammar types, I would be most delighted.

I would even agonize with them because of the idea of doing a frequency count of sentences or paragraphs is difficult not only mechanically but conceptually.

Some of my own work on sentence types showing that certain types of sentences such as those with little "distance" between the subject and verb are easier to read and remember than sentences with greater distance between the subject and the verb tends to show that certain classes of sentences operate in the reading process differently. Hence I could postulate that some sentence types could be counted, but as for objectively counting exact sentences, the prospect boggles the mind if not the largest computer memories.

There is little doubt that frequency of graphic units has a demonstrable effect on the reading process. For example, Finn (1977) found that "cloze easiness" (chance of getting a cloze item correct) correlated with frequency (Carrolls SFI) at .55. In another study, Steig (1980) using a technique similar to Shannon's Guessing Game as modified by Coleman and Miller (1968) of having students successively guess 100 words of text, found that high frequency words were guessed correctly significantly more often than low frequency words.

There is also little doubt that frequency does *not* play a prominent part in most theoretical reading models. Singer and Ruddell indexed some 484 content items in their *Theoretical Models and Process in Reading, 2nd Edition,* and "frequency" was not among them. This is despite the fact that some of the largest and, in my opinion some of the most basic studies related to reading have been in the area of frequency. Perhaps writers for a 3rd edition will not let this omission continue.

"Frequency can tell us something important about the readers use of different language units." While this is only a part of a reading theory, any reading theory which omits any consideration of frequency of unit sizes is in itself a partial theory.

REFERENCES

Ayres, L. P. *A measuring scale for ability in spelling.* New York: Russell Sage Foundation, 1915.

Carroll, J. B., Davies, P., & Richman, B. *American heritage word frequency book.* Boston: Houghton Mifflin, 1972.

Chall, J. *Learning to read: the great debate.* New York: McGraw-Hill, 1967.

Coleman, E. B. & Miller, G. R. A measure of information gained during prose learning. *Reading Research Quarterly,* 1968, *3,* 369-386.

DeRocher, J. E., Miron, M. S., Patten, S. M., & Pratt, C. C. *The counting of words.* Syracuse, New York: Syracuse University Research Corporation, 1973.

Dewey, G. *Relative frequency of English speech sounds.* Cambridge: Harvard University Press, 1923.

Dewey, G. *Relative frequency of English speech sounds.* Cambridge: Harvard University Press, 1950.

Dolch, E. W. Sight syllables versus letter phonics. *Elementary School Journal,* 1940, *41,* 38-42.

Gleitman, L. R. & Rosin, P. Teaching reading by the use of a syllabary. *Reading Research Quarterly,* Summer 1973. *8,* 447-483.

Finn, P. J. Word frequency, information theory and cloze performance: a transfer feature theory of processing in reading. *Reading Research Quarterly,* 1978, *13,* 508-537.

Hopkins, C. J. & Moe, A. J. *The computer assisted identification of common word strings from the test of children's trade books in primary grades.* Paper presented at the meeting of the International Reading Association, Houston, May, 1978.

Kucera, H. & Francis, W. N. *Computational analysis of present-day American English.* Providence, R.I.: Brown University Press, 1967.

LaBerge, D. & Samuels, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology,* 1974, *6,*293-323.

Mathews, M. M. *Teaching to read: historically considered*. Chicago: University of Chicago Press, 1966.

Osburn, W. J. Teaching spelling by teaching syllables and root words. *Elementary School Journal,* 1954, *55,* 32-41.

Sakiey, E. H. *Syllables: a weighted graphemic inventory*. Unpublished dissertation, Rutgers University, New Brunswick, New Jersey, 1977.

Steig, J. *Schema theory, word frequency and anticipation in reading*. Unpublished dissertation, Rutgers University, New Brunswick, New Jersey, 1980.

Thorndike, E. L. *The teacher's word book of 10,000 words*. New York: Teachers College Press, Columbia University, 1921.

Venezsky, R. L. & Calfee, R. C. The reading competency model. In Harry Singer and Robert Ruddell (Eds.) *Theoretical models and processes of reading*. Newark, Delaware: International Reading Association, 1970, 273-291.

ELLEN BOUCHARD RYAN
University of Notre Dame

METHODOLOGICAL PROBLEMS IN RESEARCHING STRATEGY DEFICIENCIES AMONG POOR READERS: SOME RECOMMENDATIONS

Whereas many children gradually acquire the effective sentence and passage processing skills necessary for good reading comprehension, poor readers tend to remain relatively deficient in the use of active, organizing strategies (see Golinkoff, 1976; Ryan, 1981). Much of the literature contrasting good and poor readers simply documents performance differences without providing evidence for specific strategy deficiencies and without adequate diagnosis of successful as well as unsuccessful strategies. A careful distinction is necessary between how a poor reader does perform and how he can perform. In other words, one cannot automatically infer inability from poor performance. For example, poor readers tend to ignore sentence cues to word identification, but they can use them under specific instructions or with materials highlighting their salience. Before conclusions of competence can be drawn, the effects of such manipulations of instruction and materials must be assessed.

Researchers interested in the causes of poor reading comprehension are urged to consider several methodological suggestions which promise to help clarify the distinction between competence and performance. First, separate measures of strategy use or processing activity add considerably to our interpretations of uninstructed performance as well as of the effects of instructional and materials manipulations. Belmont and Butterfield (1977) and Butterfield, Siladi, & Belmont (1980) insist that separate simultaneous measures of strategy usage should be incorporated into all comparisons between immature and mature learners to complement the performance measures. Although the emphasis among reading researchers has been on inferring strategy differences from group behavioral differences in overall scores associated with various task modifications, some direct measures of reading processes have been employed: eye movements, oral reading (intonation, errors, and corrections), eye-voice span, subvocalization, underlining, notetaking, and reading time. However, these strategy measures are rarely obtained along with an overall performance measure. All too often, reading studies have examined process without performance measures or performance without process measures (see Ryan, 1981). Furthermore, good and poor readers might differ on both process and performance measures without the performance differences resulting from the process differences. Assessments of the correlational relationships as well as group differences in both process and performance under varied circumstances can add substantially to our understanding of the nature of both successful and unsuccessful reading activities.

A second methodological suggestion is to use cognitive functional analysis and diagnosis based on the Soviet notion of the zone of potential development (Brown & French, in press; Vygotsky, 1962), an alternative to between-group designs. This individualized approach (described in the context of reading by Meichenbaum, 1976) involves a detailed analysis of the target task followed by an analysis of the performance effects resulting from systematic manipulation of task demands. The goal is to ascertain for an individual subject the nature of his cognitive strategies and the manner in which they affect performance. For any given task, the researcher develops a preliminary characterization of the required strategies by logically analyzing task demands and by observing others doing the task. However, the sequence of psychological processes involved for a given individual can only be fully appreciated after the task demands have been systematically manipulated. The presence or absence of a particular cognitive strategy can be inferred from changes in performance associated with specific task alterations. For an analysis of reading strategies, modifications in the following aspects of the task might be included: instructions, degree of tutorial assistance, difficulty of material (in terms of vocabulary, syntax, imposed organization), modality (listening vs. reading), required response (e.g., recognition, rote recall or gist recall), time constraints, and task setting. Within this paradigm, the emphasis is upon individual poor readers, not on a general description of the heterogeneous group of poor readers. Moreover, the focus is not on the failure of the poor reader but on the condition under which performance is successful and why it is not successful otherwise.

Third, training studies offer an invaluable opportunity to assess specific effects of particular strategies. For example, if poor readers can be brought up to par with their more successful peers on a comprehension task by relatively brief instruction on the use of a strategy and if the comprehension improvement is correlated with use of the instructed strategy, one can conclude that the original group difference observed is related to failure to use the target strategy and that the poor readers possess the requisite competence to succeed. If group differences are unchanged across instructed and uninstructed conditions, no support for the hypothesized strategy deficiency is obtained. If group differences persist, but are diminished under instructions, then spontaneous strategy use is implicated but cannot be taken as the full explanation for the original poor reader disadvantage. Further work would be necessary to determine whether the appropriate strategy was sufficiently well instructed; whether good readers might be more proficient in use of the strategy; whether the critical strategy for success on the task has been correctly identified; or whether the two groups might indeed differ in some aspect of basic competence.

ILLUSTRATION OF THE ANALYSIS OF STRATEGY DEFICIENCIES

The application of these recommendations can be illustrated with recent research concerning semantic integration strategies in a pseudo-reading context (see Ledger & Ryan, in press; Ryan & Ledger, in press). The Pictograph Sentence Memory Task has been specifically developed for investigating the uninstructed and trained use of deliberate sentence processing strategies. After learning a set of pictographs (picture-word symbols), children are asked to name the pictographs in a sequence and then to recall them after a brief delay. Since each sequence forms a meaningful sentence, attention to its grammatical structure and its integrated meaning facilitates the recall task substantially. Yet, young school-age children do not tend to exhibit such attention without special instructions. The sequences are presented without articles (e. g., 'girl' 'drop' 'block' 'in' 'box') so that sentence insertions (appropriate insertions of articles and verb inflections) can be used as overt reflections of sentence processing (i.e., 'The girl drops the block in the box.'). Furthermore, toys are available for acting out the meaning of each sentence so that enactions of the integrated sentence meaning or actions involving some of the separate words can be scored as further indices of the degree of sentence processing. Since deep processing can occur without actually using the toys, looking at the toys has also been scored as a potentially useful strategy. Finally, overt auditory rehearsal (the simplest useful strategy on this task) has also been recorded. Hence, analyses of performance on this task can be interpreted in terms of four types of concurrent strategy measures: sentence insertions, toy enactions, looking at toys, and overt auditory rehearsal.

In a series of four projects with children in kindergarten and first grade, we have examined the extent to which children process the sequences as sentences with and without strategy training. With neither instruction nor practice, use of sentence insertions and auditory rehearsal strategies is correlated with recall. However, virtually no children have exhibited the use of toy enactions, and looking at toys does not predict recall until an enaction or imagery strategy is implemented.

For semantic integration training, four types of sentence strategies have been examined: enaction, imagery, sentence repetition, and sentence. All four of these strategies involve the instruction to read the sequence as a sentence, inserting the appropriate articles. The most effective strategy, sentence enaction, also involves acting out the sentence meaning with the toys. Although the strategy involving only the 'read as a sentence' component significantly enhances sentence recall, the addition of the toy enaction component provides a significant advantage for children in both grades. Moreoever, training on sentence enaction eliminates the differences between good and poor recallers. Whereas the sentence reading component transfers reasonably well to similar tasks, use of the toy enaction strategy drops substantially with slight alterations of the target task. Hence, our later studies have included training on imagery strategies which involve imagining the toy enactions. Mental imagery underlies much of normal reading comprehension and its use is not tied to the presence of appropriate props. Therefore, we believed that, if young children could learn an imagery strategy, they would be able to generalize its use more successfully than the enaction strategy. Although the developmental memory literature (Rohwer, 1973) has shown that children in the first grades of school cannot yet use elaborative imagery strategies to facilitate recall of paired-associates, we have obtained encouraging evidence for the effective use of imagery by kindergarteners to facilitate deliberate sentence processing on both the target and transfer tasks. In our task, imagery involves deep semantic processing of a sentence's meaning but does not require the elaboration necessary for the paired-associate task.

In the pictograph sentence task, we have been able to utilize the three methodologies recommended for examination of the nature of poor reading. First, correlational analyses between recall and the four strategies have revealed the extent to which the children rely on each of them for their recall of the pictograph sequences both without instruction and after training. Analyses of individual differences within experimental group provide several types of evidence concerning the links between process and performance. Success on maintenance and generalization tasks tends to be correlated with the use of the components of the trained strategies. The main exceptions to this pattern have been due to ceiling effects in strategy use and to the lack of generalization of the enaction component. Also, within practice or rote repetition control groups, gradual improvement across sessions has been correlated with the uninstructed development of a sentence orientation, as measured by sentence insertions.

Training in sentence processing strategies has allowed us to clarify the causal relations between strategy and performance as well as the competence/performance issue. For example, sentence enaction training has eliminated original good/poor recaller differences as well as differences between kindergarteners and first graders. Such a finding establishes that both poor recallers and younger children have the competence required to perform the task as well as the better recallers. Also, the fact that the sentence component alone is helpful but not sufficient for both grades confirms the young child's need for additional instructions concerning deep processing of the meaning of the sentences 'read'. The post-training correlations between recall and strategy suggest that further training for some children would probably pay off with additional improvement in performance. Likewise, the gradual increase across sessions in the use of the more complex strategy, imagery, contrasts with optimal use of the other strategies in the first training session and indicates that the imagery strategy, being more difficult for young children, would require additional training to produce optimal recall levels.

In terms of cognitive functional analysis, the pictograph sentence task provides an ideal context for exploring an individual learner's needs for instructional assistance. For example, when we told three below-average second grade readers that they could enhance their performance by using the available toys, two children evidenced no change either in use of toys or in recall. The third child, however, proceeded through a trial-and-error sequence of strategies from using the toys as markers (e.g., red truck for 'red' and green block for 'block') to a partial integration (e.g., red block), and leading eventually to 'invention' of the full enaction strategy and subsequent optimal recall. A sequence of ten hints increasing in specificity could be offered sequentially to children as they attempted to work out how to recall the pictograph sequences perfectly. The more the children participate in the development of the strategy the better they should maintain and generalize its use. Furthermore, individual differences in the number of hints needed could be used to predict their subsequent performance on related tasks. With poor readers, one can begin with the pictograph reading task and move to similar tasks presented orally and with printed words. Some poor readers will probably be able to perform much better with the pictographs than the printed words, while others will exhibit little sensitivity to the change in mode. Further, the shift between oral and pictograph presentation would yield evidence concerning the child's tendency to display his sentence processing abilities better in one modality than another.

In summary, inclusion within investigations of poor reading of these recommended procedures can lead to more precise identifications of strategy deficits typical of all or some subgroups of poor readers. The Pictograph Sentence Memory task was presented as an example of the type of reading task within which strategy use can be observed and manipulated.

ACKNOWLEDGEMENT

Preparation of this article was supported in part by grants from the National Science Foundation (BNS76-09559) and the National Institute of Education (G-79-0134).

REFERENCES

Belmont, J. M. & Butterfield, E. C. The instructional approach to developmental cognitive research. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Brown, A. L. & French, L. A. The zone of potential development: Implications for intelligence testing in the year 2000. *Intelligence,* in press.

Butterfield, E. C., Siladi, D., & Belmont, J. M. Validating theories of intelligence. In H. Reese & L. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 15). New York: Academic Press, 1980.

Golinkoff, R. M. A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly,* 1976, *11,* 623-659.

Ledger, G. W. & Ryan, E. B. The effects of semantic integration training on recall for pictograph sequences. *Journal of Experimental Child Psychology,* in press.

Meichenbaum, D. Cognitive factors as determinants of learning disabilities: A cognitive functional approach. In R. Knights & D. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches.* Baltimore, Md.: University Park Press, 1976.

Rohwer, W. D., Jr. Elaborating and learning in childhood and adolescence. In H. W. Reese (Eds.), *Advances in child development and behavior* (Vol. 8). New York: Academic Press, 1973.

Ryan, E. B. Identifying and remediating failures in reading comprehension: Toward an instructional approach for poor comprehenders. In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 3). New York: Academic Press, 1981.

Ryan, E. B. & Ledger, G. W. Assessing sentence processing skills in prereaders. In B. A. Hutson (Eds.), *Advances in reading/language research* (Vol. 1). Greenwich, Conn.: J. A. I. Press, in press.

Vygotsky, L. S. *Thought and language.* Cambridge, Mass.: M. I. T. Press, 1962.

GERALD G. DUFFY
Michigan State University

TEACHER EFFECTIVENESS RESEARCH: IMPLICATIONS FOR THE READING PROFESSION

In an old *New Yorker* cartoon, a psychiatrist says to a reading teacher on the couch, "Now, as I understand it, your problem began when little Susie Jones said that she didn't have a learning disability but that you had a teaching disability."

Reading professionals have been keenly aware of the recent explosion of knowledge in language and language acquisition. The resultant focus on the interactive nature of reading and on the child's language processing has altered forever our thinking about the nature of the reading process.

However, reading is not the only research revolution in town. Another one—a revolution focused on the teaching process—has altered our understanding of life in classrooms just as the revolution in reading has altered our understanding of the reading process. Both the former, symbolized by but not limited to the Institute for Research on Teaching at Michigan State University, and the latter, symbolized by but not limited to the Center for the Study of Reading at the University of Illinois, have great potential for improving reading instruction in our schools. To date, however, the two research revolutions have been moving forward "out of earshot" of each other. Research on reading reflects little of what has been discovered about teaching; the research on teaching reflects few of the findings about the reading process.

I believe that collaborative efforts by the two research communities would have greater impact at the classroom level than independent ones. Consequently, this review of teacher effectiveness research is designed to be a first step in a dialogue between the two research communities. Since it *is* a review of teacher effectiveness, it must necessarily emphasize how research on teaching contributes to reading more than the reverse. Nevertheless, it is based on my faith that the two research communities have the potential to reciprocally inform each other and, ultimately, to make significant collaborative contributions to classroom reading instruction in the decade ahead.

BACKGROUND

The research on reading has helped us understand the implicit contractual agreement between the writer and the reader in which the writer must establish "points of contact" between the message and the reader's experience (Tierney and LaZansky, 1980). Since my purpose is to encourage a collaborative joining of two diverse research communities, it is appropriate that I provide such "points of contact" and encourage the "schema-match" necessary for the goal of this paper to be achieved. These "points of contact" are discussed in three categories: the history of teacher effectiveness research, the evolving concept of teacher effectiveness and conflicting conceptions of teaching.

HISTORICAL DEVELOPMENT

The first "point of contact" involves an understanding of the massive evolution which has occurred in research on teaching in the past ten years. The history of this evolution is presented here in two sections: the "old" and the "new" research on teacher effectiveness.

THE "OLD" TEACHER EFFECTIVENESS RESEARCH

Teacher effectiveness research is not new. To the contrary, it has been going on in one form or another for the entire twentieth century. Its impact prior to 1975 was meager, however. The AERA Committee on the Criteria of Teacher Effectiveness (Barr, *et. al*, 1952) stated the case clearly when they said:

> The simple fact of the matter is that, after forty years of research on teacher effectiveness during which a vast number of studies have been carried out, one can point to few outcomes that a superintendent of schools can safely employ in hiring a teacher or granting him tenure, that an agency can employ in certifying teachers, or that a teacher education faculty can employ in planning or improving teacher education programs.

Even twenty-two years later, Dunkin and Biddle's (1974) statement that theirs is ". . . the first text yet written that concerns the study of teaching" is a sad commentary on the significance of the "old" teacher effectiveness research.

What was the nature of this "old" research? Medley (1977a; 1979) and Duncan and Biddle (1974) provide excellent histories. Essentially, the "old" teacher effectiveness research had three thrusts.

The first focused on teacher personality and educational characteristics as they related to student attitudes and achievement. The earliest studies of this type asked pupils to describe effective teachers, with lists of traits being produced which supposedly described effective teachers. This approach failed not only because the correlations between characteristics and achievement were low and inconsistent from study to study but because, as Medley (1977a) pointed out, the research was based on:

> . . . teachers *perceived* as effective. In no instance was any evidence adduced to show that teachers possessing these characteristics were actually more effective in promoting pupil achievement of any of the goals of education. That is, the lists were never shown to be valid. Insofar as the lists described anyone, they described the teacher who *looks* effective.

Perhaps Artley (1969) provided the epitaph for this research, from the standpoint of reading professionals as well as the profession as a whole, when he said:

> As a result, they have described for us a kind of invisible, ghost-like person who, in fact, may not exist. She (he) has been found to be cooperative, sympathetic, poised. She is well-groomed, healthy, imaginative, and cooperative. She gets along with her co-workers and her principal and she gets her reports in on time. As one of my friends said, 'she has the same characteristics we expect from a good bar girl.'

A second type of "old" effectiveness research, quite familiar to the reading professional, attempted to prove one method superior to another. This research, like most "old" studies, ignored the instructional processes and classroom interactions which are the essence of teaching and, like the First Grade Studies (Bond and Dykstra, 1967) and others, contributed little to our understanding of what makes an effective teacher.

The third major type, emerging in the late 1950's and early 1960's, emphasized systematic observation of classroom interactions, particularly verbal interactions (Flanders, 1970; Medley and Mitzel, 1963). Amazingly, this is the first time that researchers began looking inside classrooms and, as such, it heralded the approach of the "new" research on teacher effectiveness which was to follow in the 1970's. Generally speaking, however, the early forms of this research had limited initial impact, largely because of what Dunkin and Biddle (1974) call "commitments"—the conducting of research in order to prove the importance of particular values held by the researcher.

In sum, the "old" research did little to explain teacher effectiveness. In 1975, Rupley (1975) was forced to describe teacher effectiveness in reading as a "puzzle phenomenon" consisting of bits and pieces which could not be fitted together to say anything meaningful about how to conduct an effective reading program. Dunkin & Biddle (1974) summarize the impact of the "old" research by saying:

> Much research on teacher effectiveness has been conducted, but little knowledge had been developed from the effort. Indeed, so little has been known that some educators have concluded that teaching is unknowable—one of the mystic skills akin to preaching or artistic creativity. Other educators have considered the activities of teaching to be so obvious as to need no research at all. Thus, master teachers write books in which they recommend the strategies they have found workable in the classroom, novelists such as Jonathan Kozol or Bell Kaufman entertain us with archetypical descriptions of teaching practices from a single school, and curriculum innovators outline new teaching programs in expectation that they will be conducted in the manner specified and that they have the effects desired. Still other educators have, in a sense, given up on teaching and concluded that pupils will learn regardless of—or in spite of—teacher's effort or that good teaching consists merely of the provision for a 'supportive learning environment.'

THE "NEW" TEACHER EFFECTIVENESS RESEARCH

The pre-1970 research did not provide substantive data regarding the nature of teacher effectiveness. What happened in the 1970's to change this?

The critical difference was that observations came into use to determine what goes on inside classrooms. Observation was not a regular part of the "old" research studies and no one knew what was actually happening, particularly as it related to the instructional interaction between teachers and students. The result was that *assumptions* about what happened in classroom instruction went largely unchallenged and teaching principles, generalizations and theories based on these assumptions were uncritically accepted. Recent research, in contrast, has been dominated by classroom observation. Two methodologies have emerged in which observation has been used in different ways to achieve different purposes: process-product studies and descriptive studies.

Process-product research. Process-product studies employ observation tools which trained observers use when visiting classrooms to record how often a particular phenomenon noted on the observation form occurs in actual practice. The combined observations are analyzed, with the individual teacher as the unit of analysis, to determine the correlation between particular coded items and achievement growth as determined by standardized achievement tests and/or by less formal attitude measures. Influenced by the behavioristic tradition, the focus has been the overt acts of teachers and the relation between the frequency of these acts and various measurable outcomes. Historically, such research had not been very successful. However, Rosenshine and Furst (1973) and Dunkin and Biddle (1974) created a comprehensible mosaic out of what had previously been a confused and largely contradictory mass of data and are credited with providing the springboard for the growth and acceptance of this type of research. Subsequent methodological improvements (see Brophy, 1979a) have since solidified its position.

The process-product findings, expressed as middle range correlations which have been demonstrated to be consistent from study to study, point to a strong relationship between certain clusters of teacher behaviors and student outcomes. Further, and most encouraging, these constellations of behaviors have been validated in experimental studies (for excellent examples of such experimental validations of process-product findings, see Anderson, Evertson and Brophy, 1979; Good and Grouws, 1979; and Stanford's Program on Teaching Effectiveness, 1978). Medley (1977a) summarizes the impact of the process-product research by stating that the ". . . risk that any of them [findings] will be disproven by future research is slight."

Nevertheless, process-product research has not been without critics. Conceptually, it has been criticized as too narrow and prescriptive (Cazden, 1977; Doyle, 1977), too simplified in the face of classroom complexities (Fisher and Beliner, 1977) and too behavioristic rather than cognitive (Lanier and Shulman, 1978). Methodologically, the potential lack of reliability and validity in the observation measures, and the design/analysis weaknesses associated with correlational studies have all been criticized (Koehler, 1977).

Descriptive research. Such criticisms contributed to the rise of descriptive, or qualitative, research. As Koehler (1977) has pointed out, qualitative research attempts to describe or define the process; to determine what *is* rather than what should be. It is less concerned with improvement and more concerned with understanding. Despite the common devotion to classroom observation, the differences between descriptive and the more improvement-oriented process-product research are startling. While the process-product researcher observes specific behaviors with a predetermined observation form in multiple classroom settings, the descriptive researcher observes a limited number of classrooms (sometimes only one), avoids specifying the precise behaviors being sought and records the classroom life (as well as impressions of this life) in free-form field notes. In contrast to the process-product research, descriptive research comes from an anthropological tradition, is heavily influenced by cognitive psychology and cognitive information processing and serves as an umbrella for a variety of types, including ethnography, participant observation and sociolinguistics, among others. Like the process-product research, it has received criticism, with the focus being on the general absence of outcome measures, the reliability problem inherent in qualitative observation, limited sample size and problems of generalizability. Despite criticism and diversity of style, however, descriptive research has produced a rich array of findings in the past five years, tempering tendencies to use process-product findings in narrow, prescriptive ways by highlighting the variability of behavior in the classroom, the multiple complexities of classroom life, the immense information processing task faced by teachers and the multiple sociological forces which interact in the classroom.

Summary

The historical evolution of research on teacher effectiveness is the first "point of contact" between the reading and teaching communities. Because of the shift to observation of classroom processes and because process-product and descriptive designs are currently being used in creative combinations, the nature of teacher effectiveness research has changed dramatically in the past ten years. To chalk off research on teaching on the basis of "the way it was" would be an error.

THE EXPANDING CONCEPT OF TEACHER EFFECTIVENESS

The above historical evolution has expanded the concept of teacher effectiveness, encompassing subtleties and complexities not traditionally associated with the term. This modification is a second "point of contact".

Prior to about 1975, teacher effectiveness research focused on determining the extent to which certain teacher behaviors produce greater amounts of specified pupil outcome. By 1975, this research, largely conducted within the process-product paradigm, had determined that certain patterns of teacher instructional behavior make a major difference in pupil achievement.

This finding spurred a second research thrust, beginning in about 1975, in which the importance of the teacher was assumed and the questions focused more on *why* some teachers were more effective than others. The basic hypothesis was that instructional information processing and instructional decision-making were the keys. However, the studies, primarily descriptive in nature, contradicted the hypothesis, suggesting instead that teaching is so complex that teachers must limit the amount of information they process and that, in fact, the demands of the workplace prevent them from making significant decisions during the interactive phase of teaching.

These findings have triggered a third set of studies, currently underway, which examine the complexities of the workplace as a means for explaining why teachers are less reflective than hypothesized. These studies have taken three forms, depending upon the breadth with which the workplace is viewed. All, however, are united in the belief that simply correlating teacher behavior with pupil outcomes is insufficient in itself for understanding teacher effectiveness.

At the first level, the workplace is the classroom, with the complexity of teaching being explained both in terms of technical and social demands. Fisher (1980) describes the technical demands:

> Some of the broad areas in which teachers must be competent include organization, management, interpersonal relationships, planning and evaluation . . . Within these broad areas, teachers must develop and refine a multitude of sub-skills to accomplish the diagnosis, prescriptions, the presentation, monitoring and feedback functions which support day-to-day

> learning in the classroom. These skills cannot be implemented one at a time. The nature of teaching is such that several skills are required simultaneously.

Shulman (1980) describes the case for the social complexity of classrooms:

> Classrooms are complexly nested social systems in which participants are engaged in jointly produced educational activities, reciprocally caused and effected, negotiating for opportunities and competing for attention and understanding.

Hence, at the classroom level, research focuses on both technical and social complexities. The explicit curriculum, the "hidden" curriculum of implicit rules and procedures which govern classroom life and the ecological relationships among them are all examined in order to understand why teachers appear to react in conditioned ways rather than as reflective decision-makers.

At the next level, the workplace involves more than the classroom. Here, Shulman (1980) argues that the teacher's world encompasses the multiple role demands of the tutor, the classroom teacher, the curriculum planner and the organizational member. He suggests that the information processing of teachers is explained, at least in part, by a "role strain" resulting from the teacher's attempts to accommodate the multiple demands of these four contexts. As such, he sees teaching as more than information processing; it is also social role processing and we cannot understand the teacher's mental life and its impact on effectiveness unless we first understand the demands of the multiplicity of teaching contexts and their respective task environments. As he summarizes:

> The teacher is working simultaneously as a tutor, instructor, pedagogue and organization member. Thus, the challenge of teaching is characterized by the need to work effectively within each level while also nimbly negotiating one's way across levels. Moreover, our concept of teacher effectiveness will surely be different depending upon the level at which we choose to define and assess effectiveness . . .

At its broadest, the teacher's workplace is viewed as society as a whole. As described by the "new sociologists", the complexities of teaching can be attributed to pressures which go beyond even the organization level described by Shulman. In this view, these complexities are best understood when examined from the perspective of the teacher as a mediator and negotiator of a variety of explicit and implicit pressures and mandates from the total society (Schwille, Porter, Gant, Belli, Floden, Freeman, Knappan, Kuhs & Schmidt, 1979).

Hence, the second "point of contact" between the reading and teaching communities lies with understanding that the term "teacher effectiveness" has been modified. The former relatively simple concept of a uni-directional linear process in which teacher behaviors caused student outcomes in an in-put/out-put relationship has evolved into a view of teaching as reciprocal interactions between teacher and context with effectiveness being a multi-faceted function of this interaction. This paper, while fulfilling the assigned charge of reviewing the "teacher effectiveness" research, reflects the more complicated concept which is emerging from recent research.

CONFLICTING CONCEPTIONS OF TEACHING

The third "point of contact" between reading and teaching researchers is embedded in the conception each holds of teaching. Despite past references to the importance of the teacher by reading professionals such as Bond and Dykstra (1967), Ramsey (1962), Artley (1969) and others, the reading profession as a whole has placed little emphasis on the actual act of teaching. Reading educators, for instance, are more likely to debate whether the reading curriculum should reflect holistic or sub-skills analyses or whether basal texts or language experience is the "best approach" than they are to discuss what teacher behaviors might stimulate desired reading outcomes. Further, while reseachers on teaching emphasize *teacher* mediation and guidance of learning, reading professonals emphasize *learner* mediation, arguing that teacher guidance impedes reading acquisition by constraining the "naturalness" of learning. Bringing these two divergent backgrounds together is difficult. In an attempt to do so, the differences will be discussed first as a function of situational context and, second, as a function of preserving the naturalness of language during instruction.

As a problem of situational context, the differences in conceptions of teaching can be understood by returning to Shulman's (1980) four levels of context: tutor, classroom instruction, curriculum development and organization member. Historically, reading researchers and eduators have focused on understanding the nature of the reading process (a function of the third level); when translating this knowledge to practice, the emphasis has been on describing what reading is in methods courses and methods textbooks (a third level function), organizing reading curriculum into basal texts, skill hierarchies and curriculum guides (again, a function of the third level) and graduate level preparation of clinicians and remedial reading specialists (a function of the first level). Hence, the reading researchers and educators emphasize levels one and three. In contrast, however, researchers on teaching emphasize the second level—classroom instruction. This difference in perspective accounts for much of the difference in conceptions of teaching. A common ground between the two groups might be the fact that classroom teachers represent the largest clientole for both research communities. While we in reading must certainly maintain our focus on the nature of reading, we nevertheless have a simultaneous responsibility for communicating our findings in terms of the classroom context shared by the bulk of our clientele. By assuming this responsibility, we could close the gap between the teaching conceptions held by reading and teaching researchers respectively.

Regarding the "naturalness" of language acquisition, research on the reading process suggests that comprehension is created by the reader, not the teacher, in response to the text and that comprehension is not something that teachers can place directly into the heads of pupils. On the other hand, researchers on teaching report convincing data indicating that patterns of teacher behaviors *do* make a difference in producing achievement gains (including gains in comprehension) and that many of the crucial behaviors involve techniques of direct instruction, as opposed to more passive instructional approaches which call for less teacher intervention. Herein lies a conflict. Perhaps the two groups can be brought closer together on this issue by re-considering the concept of "naturalness". Because pupils create meaning out of their personalized experience backgrounds does not necessarily mean that teachers should not expedite this process by showing them how to create meaning; it only means that the guidance should not force the child to substitute the teacher's interpretation for his/her own. Hence, teachers can preserve the naturalness of meaning-getting while still performing the professional task of teaching pupils *how* to get meaning. Collaborative efforts by the two research communities in this area could be quite fruitful.

Summary

"Schematic mis-matches" could impede collaboration by the reading and teaching research communities. Hopefully, the above discussion of the evolution of teacher effectiveness research, the resultant modification of the concept of teacher effectiveness and the potential conflict in conceptions of teaching will enhance communication.

However, it should also be noted that communication is two-edged. The writer has a responsibility, but so does the reader. As stated by Tierney and LaZansky (1980):

> . . . it is the responsibility of a reader to bring to a text conceptualizations which support a reasonable interpretation of the text rather than result in an abandonment of the author's message.

MAJOR TEACHER EFFECTIVENESS FINDINGS

The seventies were a decade of growth for research on teaching. Many studies, both process-product and descriptive, were funded by the National Institute for Education or were conducted independently. The bulk of these focused on the teaching of basic skills (mostly reading and mathematics) in school settings where academic achievement has been traditionally at or below national norms. A multiplicity of findings emerged but they can be discussed in terms of two major conclusions:

1. The most effective teachers of basic skills generate the most opportunity to learn; and
2. such teachers are technical managers of instructional materials and activities rather than theory-driven and reflective decision-makers.

These conclusions can have a far-reaching and significant impact on our perception of classroom teaching. However, as Medley (1977a) has pointed out:

> . . . the profile of the effective teacher that emerges is not entirely consistent with the general consensus of how a good teacher behaves or with the way teacher educators train teachers to act.

Two points should be made about the above lack of congruence. First, the conclusions about teacher effectiveness represent classroom life as it currently is. Second, such reality does not necessarily represent either what effectiveness should be or what it can ultimately become. The findings *do*, however, serve as an important foundation for learning more about the teaching process and for making future teachers more effective than current ones.

OPPORTUNITY TO LEARN

Opportunity to learn is generally discussed in terms of Carroll's (1963) model of school learning in which three pupil variables—aptitude, ability and perserverance—interact with both opportunity to learn and quality of instruction to account for learning outcomes. The findings from research on teaching reflect three components of the opportunity to learn variable: direct instruction, management and psychological conditions.

Direct Instruction

Direct instruction is a global concept, largely associated with reviews by Rosenshine (1976; 1979; 1980), which encompasses the variety of behaviors displayed by teachers in classrooms where learning gains are great. These behaviors can be discussed in terms of engaged time and teacher monitoring of pupil activity.

In the classroom, a subject such as reading is allocated a certain amount of instructional time. Engaged time is that proportion of the allocated time in which the student is on task or "engaged". The more time a student spends engaged in an academic endeavor, the more that will be learned; the less time engaged, the less that will be learned (Berliner, 1979; Rosenshine and Berliner, 1978). As Fisher, Berliner, Filby, Marliave, Cahen and Dishaw, (1980) said:

> . . . two classes might allocate the same amount of time to reading instruction, but one class might have almost twice as much real engaged learning time as the other. Since engagement rate has been shown to be highly variable across classes and since that variability has been empirically related to achievement, it is possible that increasing engagement rates will lead to increased achievements.

The engaged time concept is a clear and simple one. Unlike most clear and simple notions in education, however, this one repeatedly discriminates effective instruction from ineffective instruction.

The second component of direct instruction focuses on teacher monitoring behaviors which insure that pupils, once engaged, will remain engaged. These behaviors imply teacher-pupil interactions which reading professionals have traditionally demonstrated some discomfort with but which have been validated in basic skills contexts in both primary grades (Anderson, Evertson and Brophy, 1979) and in remedial secondary classrooms (Stallings, 1978). They indicate that the most effective teachers are those who are the most structured, the most in control and the most directive; they are the teachers who monitor pupils' activities closely, who call for frequent repetition and who drill; and they are the teachers who move in small steps, who teach to over-learning and who elicit response from and provide feedback to each individual student (Berliner, 1979; Rosenshine, 1979; Brophy, 1979c). Such teacher interaction is called "high structuring" by Gage (1978) while Good (1980) prefers to call it "active teaching". Regardless of its label, however, the data indicate that these behaviors do distinguish effective basic skills teachers from ineffective ones.

In summary, direct instruction creates opportunity to learn by engaging pupils and by insuring that such engagement is maintained. Perhaps Koehler's (1977) statement about direct instruction is the most descriptive:

In this model, the teacher supervises lessons and workbook activity, allows little free time or unsupervised student desk work and clearly communicates the goals of the lessons to the students. The teacher decides what activities will take place, but the students are actively involved in the lesson. Content coverage is extensive, and questions tend to be focused at a low cognitive level. These questions are geared to information which should already be known rather than to that which can be deduced or guessed. Teacher reinforcement rapidly follows most answers. Learning is organized around questions posed by the teacher or materials provided by the teacher. Teacher-student interactions are direct and academically oriented.

Classroom Management

Effective teachers are also good classroom managers; they create more opportunity to learn by organizing themselves and the classroom to enhance efficiency and minimize time wastage. The importance of management is stated succintly by Good (1980):

> . . . teachers' managerial abilities have been found to relate positively to student achievement in *every* process-product study conducted to date . . . It appears that teacher managerial skills are necessary if reasonable pupil achievement is to occur . . .

Its importance in reading particularly is emphasized by Leinhardt, Ligmond and Cooley, (1980), Duffy (1979), Roehler and Schmidt (1980), Zigmond, Leinhardt and Cooley (1980), and others who point to the time lost because of poor management.

Management is another aspect of teaching which causes reading educators some difficulty. This is so for three reasons. First, reading people seem to equate classroom management with skills management systems. This is an erroneous association. Classroom management refers to techniques for generating maximum instructional time; monitoring pupil progress through a hierarchy of skills *may* be one aspect of generating more instructional time but the two concepts are not synonymous. Second, the academic heritage of reading educators is often tied to the tutoring of remedial students where time management is less complex. In contrast to the clinic where the pupil typically receives the tutor's individual attention, time in the classroom is a limited commodity demanding what Shulman (1980) calls "distributional rules." Hence, the classroom teacher faces an "opportunity to learn" problem seldom faced by the clinician. As Brophy (1980b) says:

> So long as teachers must deal with classes of 25 or 30 instead of tutoring single individuals, there will be trade-offs between meeting the needs of any individual and meeting the needs of other individuals or of the class as a group.

Third, even when reading educators understand the purpose of management, they tend to reject it because a common trade-off of teachers is to use routines as a way to maximize instructional time and, to some people, routines imply mechanistic teacher-pupil relationships. This is unfortunate since, rather than inhibiting humanistic interactions, routines are the vehicles effective teachers use to "distribute" time so that they are personally engaged in as many meaningful interactions as possible while limiting "traffic manager" and disciplinarian interactions to a minimum.

The secret to effective management is the teacher's ability to establish such routines as preventive measures. As Good (1980) has said:

> The key behaviors that distinguish good classroom managers are those techniques which *prevent* misbehavior by eliciting student cooperation in general and involvement in assigned work specifically.

In addition, according to Anderson, Evertson and Emmer (in press), the most effective teachers convey the purposefulness and meaningfulness of academic activities, instruct students in the skills of good behavior and select activities that reflect both students' level of understanding and their need for information. Other crucial aspects of management are reported by Kounin (1970) and by Brophy and Putnam (1979).

In short, good management creates more opportunity to learn (or, more specifically, preserves for instructional purposes larger amounts of the allocated instructional time). The strong relationship between management, instruction and greater learning outcomes has caused Brophy (1979d) to suggest that management skills are closely linked with instructional skills—that good instructors tend to be good managers and vice versa. Hence, classroom management is one of the behavior clusters which distinguishes effective teachers from ineffective ones.

Psychological Conditions

Discussion of opportunity to learn is often confined to the purely physical notion of time and time usage. However, research on teaching also suggests that effective teachers create "psychological" opportunities to learn as well as physical ones. Through their instructional behaviors, they create a climate which encourages pupils to make maximum use of the physical time provided. These psychological conditions include expectations, efficacy and success experiences.

Expectancy is the teacher's perception of how much (and how quickly) a student can be expected to learn. Studies by McDonald and Elias (1976) and Brookover, *et. al.* (1978) in the United States and by Rutter, Maughan, Martimore, Ouston & Smith (1979) in the United Kingdom all indicate that effectiveness is distinguished from ineffectiveness by the teacher's belief that students can and do learn.

Efficacy is also a type of expectancy but refers to the teacher's perception of his/her ability to be successful as a teacher and the level of effort and persistance exhibited as a result. Reports by Brophy and Evertson (1976) and by Good (1980) support the importance of efficacy. As Brophy (1979d) says:

> Typically, the more successful teachers have a 'can do' attitude, perceiving their students as capable of learning the material and themselves as capable of teaching it to them effectively. These teachers set higher goals than other teachers, and they are more persistant in laboring to meet these goals and overcoming obstacles if necessary.

The third psychological condition is student success. When students are engaged in tasks in which they are successful, they achieve more than when they are engaged in tasks in which they are unsuccessful. This concept is so crucial that the Beginning Teacher Education Study, prime investigators of time as it relates to teacher effectiveness (Fisher, *et. al.*, 1980), includes success as part of their definition of engaged time. The implication, of course, is that time spent engaged in an excessively difficult task creates negative pupil behaviors which impede opportunity to learn. In addition, pupil success is tied to improved pupil attitudes. As Fisher, *et. al.*, (1980) state:

> It is interesting to note that the high success component of learning is associated with more positive student attitudes. Successful students probably enjoy learning more because of their success. Failure, even when it is only occasional, appears to result in more negative attitude . . .

Brophy (1980a) concurs, saying, "Students consistently given work that is too difficult for them can be expected to give up and eventually to become 'motivation problems'."

In summary, effective teachers apparently motivate by letting pupils know that they are capable of learning, by exuding confidence in their own capacity to provide useful assistance to pupils and by assigning tasks at which pupils can succeed, thereby making the prophecy come true. Other psychological conditions may be the teacher's acknowledgement to pupils of the difficulty of a learning task (Duffy and Sherman, 1977; Green, 1978) and the matching of student and teacher styles (Brophy, 1980b).

Summary of Opportunity to Learn Findings

The teacher effectiveness research, particularly as it relates to basic skills instruction, indicates that teacher behaviors which create opportunity to learn are crucial for producing growth in pupil achievement (and, to some extent, attitudes). Properly understood, these findings can help us improve reading instruction.

Unfortuntely, however, they are not always properly understood. For instance, it is sometimes believed that the opportunity to learn findings represent support for discrete teaching behaviors in the tradition of competency-based instruction when, in fact, the findings refute the concept of generic teaching behaviors in favor of constellations or clusters of behavior (Brophy, 1979a; Good, 1980); it is sometimes believed that these findings will encourage the mandating of prescriptive algorithms which will make the teacher mechanistic and non-humanistic when, in fact, teaching researchers such as Gage (1978), Good (1980) and Fenstermacher (1980a) argue eloquently for teacher judgment in implementing the findings; it is sometimes believed that phonics is emphasized to the detriment of comprehension when, in fact, successful comprehension programs such as the Kamehameha Project reflect opportunity to learn findings (Au, 1980; Sloat, 1980; Tharp, 1980); there is a belief that these findings apply to all reading instruction when, in fact, they apply only to basic skill situations (Brophy, 1980b; Good, 1980); there is the belief that the behaviors associated with opportunity to learn are universal truths when, in fact, effective teaching varies according to a multiplicity of contextual factors (Brophy, 1980b; Good, 1980; Peterson, 1979); and finally, there is the belief that the findings are generated exclusively by behaviorists when, in fact, research on teaching reflects an interdisciplinary balance between behaviorists, cognitive psychologists, anthropologists, sociologists, educators and others.

The message, I hope, is clear. Collaboration between reading and teaching researchers requires accurate interpretation of what the opportunity to learn findings do and do not imply.

TEACHING AS TECHNICAL BEHAVIOR

We have long felt that teachers *ought* to be reflective, we have pointed to the rebellion of teachers involved with "teacher proof" reading programs as evidence of what happens when a teacher's rationality is stifled and we have assumed that teaching is guided by reasoned rather than conditioned or reactive behavior. This belief, together with the emerging interest in cognitive psychology and cognitive information processing, has led many researchers on teaching to look beyond direct instruction, management and psychological conditions to determine the goals, intentions, judgments, decisions and information processing which undergirds the behavior associated with teacher effectiveness. In short, teacher effectiveness reseachers have begun to examine the nature of teacher rationality on the assumption that their behavior must be guided by what they think. Parallel studies in reading have hypothesized that theories and models of reading influence a teacher's instruction (Harste and Burke, 1977; Buike, Burke and Duffy, 1980; Kamil and Pearson, 1979), that clinical and remedial decision-making in reading reflects a rational model (Vinsonhaler, 1979) and that teachers can reflectively apply diagnostic-prescriptive techniques in teaching reading (Duffy, Sherman and Roehler, 1977).

Such interest in the teacher's mental life has produced two major categories of research on teacher thinking. They focus on the teacher as a planner and on the teacher as a decision maker.

The Teacher As A Planner

Clark and Yinger (1980) have reviewed virtually all the planning research. They report that teachers do not think about planning in the way we have assumed they did. Rather than following the objective-based, linear model of planning promoted in most teacher education programs, teachers apparently initiate planning by selecting an activity which they then fit to the time available and to other constraints. Only then, somewhat as an afterthought, do they consider what specific skills or processes the students may learn by pursuing this activity. Other reviewers and researchers (Brophy, 1980a; Morine-Dershimer, 1979; Joyce, 1980) also indicate that plans are activity-focused and routinized in order to expidite teacher monitoring functions and to insure ease of management. In short, the research on planning pictures teachers as technicians who manage activities rather than as professionals who reflectively select strategies and tactics to achieve particular goals.

The Teacher As A Decison-Maker

Like planning, the concept of the teacher as an interactive decision-maker reflects the ideal of rational teacher behavior. As described by Clark and Yinger (1980):

> Interactive decision making refers to decisions made during the act of teaching. The teacher is seen as constantly assessing the situation, processing information about the situation, making decisions about what to do next, guiding action on the basis of the decisions, and observing the effectiveness of the action on students.

While laboratory research appears to substantiate the concept of interactive decision-making, classroom observation studies indicate that teachers are not as reflective "on the job" as had been assumed. Instead, as noted by Joyce (1980), teachers tend to emphasize the maintainance of on-going activity flow in the classroom rather than interactive decision-making. In reading particularly, the apparent lack of decision-making during the interactive stage has led Buike (1980) and Duffy and McIntyre (1980) to characterize the teacher as a "technician" and as "activity-driven." Durkin's (1979) study of classroom comprehension instruction (or, rather, the lack thereof) further suggests the absence of substantive instructional decision-making, as does the report by Tierney and Pearson (1980) that teachers object to open-ended comprehension questions in the basal text.

Possible Causes of Technical Behavior

These findings about teacher planning and decision-making are difficult to accept when we *want* to believe that teachers are basing their actions on rational instructional models. However, they are not so surprising when examined from the standpoint of classroom realities generally and reading instruction in particular.

Let us consider classroom realities. The first reality is that failure to establish an on-going activity flow results in serious behavioral management difficulties. Consequently, teachers convert a complex and potentially explosive group of thirty children into a predictable and simplified routine which can be monitored with relative ease (Brophy, 1980b; Shavelson, 1980). Jackson (1966) summarizes the situation best when he describes the teacher's concern as:

> . . . making some kind of educated guess about what would be a beneficial activity for a student or group of students and then doing whatever is necessary to see that the participants remain involved in that activity. The teacher's goal, in other words, is student involvement rather than student learning. It is true, of course, that the teacher hopes that the environment will result in certain fundamental changes in the students. The learning is, in this sense, a by-product or a secondary goal rather than the thing about which the teacher is most directly concerned.

Second, as both Brophy (1980a) and Shavelson (1980) suggest, technical teacher behavior makes sense in terms of cognitive information processing. Interrupting an established activity flow in a classroom to reflect on and consider an alternative (while simultaneously trying to keep thirty pupils involved) increases significantly the information processing demands placed on the teacher. In effect, a "cognitive overload" occurs and the teacher must simply move on. Most of us who have taught school for any length of time admit that this, indeed, is the way it often is.

A third hypothesis, noted earlier, is that the demands of the workplace, whether viewed as the classroom, as multiple role strains or as society at large, account for the paucity of observed teacher reflection during instruction. While a different kind of decision-making may be occurring in contexts other than public lessons (Shulman, 1980), the complexities of the workplace severely limit interactive decision-making.

Finally, the fourth hypothesis reflecting classroom realities is seen in the work of Barnes, Putnam and Wanous (1979). They have created a teacher education program based on the concept that prospective teachers must be effective classroom managers before they can be effective instructors. In effect, they are saying that there must be "automaticity" in classroom management in order for teachers to make room in their information processing apparatus for interactive decision-making. Hence, teachers who must consciously attend to the technical demands of management become technicians; if these tasks become automatic, the teacher is free to concentrate on instructional decision-making.

Now let's look at how the nature of reading instruction itself helps explain why so many teachers look like technicians. First, in analyzing the research being conducted at the Institute for Research on Teaching on clinical reading diagnosis and the data which repeatedly indicate that practicing and certified diagnosticians are not even reliable with themselves much less with one another, Brophy (1980a) suggests that the fault lies with the primitive nature of our knowledge base in reading. In other words, reading teachers have difficulty being reflective because there is not enough reading knowledge to allow them to be specific and truly diagnostic. Hence, they must be technicians. As Brophy (1980a) says:

> In educational diagnosis and remediation, it is as if we are treating all patients by telling them to get some rest, take aspirin and drink a lot of fluids no matter what their problem . . . Until we develop a knowledge-base to support more truly diagnostic and remedial procedures, I do not think that we can expect much from even 'experts' let alone ordinary classroom teachers.

Second, reading teachers may behave like technicians because of the prominent role played by materials, particularly the basal textbook (Clark and Yinger, 1980; Osborn, 1980). Teachers are encouraged and often mandated to use basals, but these materials themselves are designed for use by technicians. The teacher's guide, as recently noted by Durkin (Richey, 1980), suggests many more practice activities than instructional suggestions and the texts, workbooks and other materials are all designed to be used in a question-and-answer, recitation format. As my colleagues and I have been discovering in our research on teacher conceptions of reading (Buike, Burke, and Duffy, 1980), the basal materials rather than theoretical considerations, rational models or reflective procedures dominate teacher thinking because basals aid in resolving management concerns while simultaneously being professionally acceptable.

Third, teachers behave as technicians because we inadvertently promote such behavior. It is not unusual, for instance, for the teacher to be described as a mediator of materials (Gage, 1978); for instruction to be characterized as technology (Barr and Dreban, 1977); for teachers to be seen as little more than custodians of curriculum (Coltheart, Note 1); for methods texts and sessions at professional reading conferences to emphasize activities, games and drills designed for recitation formats; for methods courses to promote such patterned teacher monitoring behavior as SQ3R and the Directed Reading Lesson, and for good comprehension instruction to be equated with "asking the right questions" which, as Durkin (1979) points out, is often just an assessment technique. Even prominent reading professionals such as Barr (1973-74) and Durkin (Richey, 1980) tend to equate instruction with the essentially technical tasks of instructional pacing and use of teacher's guides.

Finally, reading teachers may become technicians in self-defense. Not only do we supply them with little other than technical "tricks of the trade," we often give the impression that reluctant readers can be transformed into fluent readers simply by caring about them, by letting them "learn to read by reading" or by letting reading acquisition happen naturally. Good (1980) addresses this issue when he suggests that our teacher education programs may contribute to low teacher efficacy by giving prospective and in-service teachers unrealistic expectations regarding what to expect when teaching reading. As he says:

> Teachers need to understand that teaching is a very tough but feasible task. Teachers' expectations for simple solutions to problems can turn into anger, frustration and, ultimately, withdrawal in the face of classroom realities.

In short, teachers may be technicians because we create them. We tell them that reading acquisition is a natural phenomenon requiring little teacher effort or guidance and we provide instructional suggestions which limit their actions to that of monitoring and reacting rather than presenting, stimulating and guiding. When teachers have been prepared in this way and are then faced with the realities of classroom life, what recourse do they have but to become technicians?

Summary of Technical Behavior

In sum, the above findings contradict some of our most cherished beliefs about teachers' reflective behavior. Rather than being driven by goal-oriented and theoretically-consistent instructional models, teachers appear to be preoccupied with activities which maintain activity flow. Consequently, teachers look like technicians and managers rather than like reflective professionals.

There are, however, three points which temper this conclusion. The first, as noted above, is that we may well be contributing to the creation of such technical behavior and we must be prepared to do more than simply condemn it.

Second, it must be noted that the generally unpalatable picture of the teacher as a technician is nevertheless real. These data help us to remember that the teacher is a normal person who, despite frailties and limited resources, must cope with that fluid, complex, demanding and potentially volatile place called a classroom. While the findings may be alien in terms of a reading researcher's schema, they are no less valid for that and we must deal effectively with this reality if we are to communicate with our clientele.

Finally, Good (1980), Fenstermacher (1980b) and McDonald (1977) suggest that the ultimate key to true teacher effectiveness may lie with teacher judgment in rationally applying suggested findings rather than in unquestioningly applying algorithms and prescriptions. While good technicians may be more effective than poor technicians, the ultimate in effective teaching will require more than technically competent teachers.

SUMMARY OF TEACHER EFFECTIVENESS FINDINGS

In [illegible]ne ways, it is overly simple to summarize the past decade of research on teaching in just two conclusions about opportunity to learn and teaching as technical behavior. However, these conclusions capture the essence of the reality of classroom life, and it is this reality which is the major contribution of research on teaching. With these conclusions as a base, we can hopefully move to studies which not only describe what is currently "effective teaching" but also what effective teaching can ultimately become.

IMPLICATIONS FOR THE READING PROFESSION

The teacher effectiveness research places reading education squarely between assumptions about "what ought to be" and the reality of "what is." On the one hand, we argue that reading acquisition *ought to be* a personalized interaction between child and text which develops with little artificial intervention from teachers, and that teachers should function in indirect and unstructured ways because they are reflective professionals. In contrast, the research on teaching suggests that effective teachers of reading (especially beginning reading) generate opportunity to learn by establishing efficient management, structured learning, and psychological conditions and that they keep their sanity by deliberately minimizing the need for "in-flight" decisions.

The conflict between "what ought to be" and "what is" is a painful one. If left unresolved, it can weaken chances for collaborative work involving reading and teaching researchers. Since the goal of this paper is to build collaboration, the following implications are presented in terms of how research on teaching can help us shape our reading education efforts, influence our reading research and provide collaborative agendas for the future.

Teacher Education

The first implication of research on teaching is that communication with prospective and in-service teachers must include acknowledgement of the realities of classroom life. To prepare classroom teachers without giving adequate attention to direct instruction, management, psychological conditions and the current limitations in the classroom teacher's information processing capacities results in a distorted picture of what the world of the classroom is really like. Ultimately, a backlash occurs; the schema of the reading educator conflicts so dramatically with the practitioner's schema that the entire message about reading is rejected. If our goal is to have teachers apply the emerging findings about the reading process, we must present these findings within "points of contact" representing the experiences teachers have had in the real world of classrooms. Failure to do this risks the rejection by teachers of all the potential benefits of reading research.

Second, we must examine the teacher effectiveness findings not only for the clues they provide regarding contextually-valid vehicles in which we can embed our messages about the nature of reading but also for their potential in helping teachers become better instructors of reading. In short, we must help teachers not only with the nature of what is taught (at Shulman's third level) but with *how* it is best taught (the first and second levels). This means that we must deal with the pedogogical problems inherent in equitably distributing opportunity to learn among thirty children and with how teachers are to find room in their already overcrowded cognitive world for the information processing demands inherent in the instructional recommendations we make.

Third, we must highlight the limitations, as well as the strengths, of the instructional findings. A major limitation is that effective teaching in one context is not necessarily effective teaching in another. Consequently, reading educators must help teachers exercise judgment in applying teacher effectiveness findings to appropriate contexts and should avoid promoting universal theories and models of reading instruction which imply the use of identical instruction in all contexts.

Finally, we must recognize that the findings on teaching present a *partial* picture of instruction. The most glaring example of this is the preoccupation with Carroll's concept of "opportunity to learn" and the relatively minor emphasis on "quality of instruction." In effect, the weight of the quantitative time usage data has caused researchers of teaching to equate quality with maximized opportunity to learn. Virtually no research has been conducted which examines effects of the qualitative guidance or assistance teachers offer as a prelude to or as part of the structured direct instruction activities nor has anyone examined the ways opportunity to learn might be modified as a result of teacher guidance. I personally suspect that one reason why so many teachers conduct mechanical recitation sessions is that they have no concept of how to qualitatively assist pupils in their learning. They only know how to keep them on task. As Good and Brophy (1978) state:

> . . . teachers sometimes act as if the students are supposed to learn on their own with no help from them . . . Such behavior represents a fundamental failure to appreciate the teacher's basic role. The teacher is in the classroom to instruct.

Instruction is more than creating opportunity to learn. Research on teaching, unfortunately, provides little specific assistance regarding this additional dimension of teaching.

Reading Research

Teacher effectiveness findings suggest four implications for reading research. First, reading research which focuses on "instruction" should, at the minimum, demonstrate an awareness of classroom realities and control for crucial teaching variables. Conducting research on the teaching of inferencing without controlling engaged time is one example of how ignoring findings on teaching can taint research on reading; to compare teachers of reading without attempting to account for efficacy is another; and conducting reading research in laboratory contexts or with college students and suggesting how the findings should be applied in elementary classrooms is a third. Researchers of reading, especially those studying comprehension, have an obligation to either control for the realities of instruction in the design of the study or to acknowledge its absence when reporting the data.

Secondly, reading researchers, like teacher effectiveness researchers, must begin looking more systematically at various contexts. The results of research on teaching speak clearly to the point that what is true at one grade level, with one ability group, in one SES setting is not necessarily true for another. Sweeping generalizations cannot be made about teaching generally and they cannot be made for reading instruction in particular. Consequently, conducting reading research in multiple contexts is essential.

Third, reading researchers should consider making greater use of qualitative research designs. To date, descriptive research has been little used by reading researchers, particularly those studying reading comprehension instruction. However, a program of such research would go a long way in informing comprehension researchers of the complexities of classroom instruction, of the limitations of teachers' information processing capacities "under fire" and the realities within which laboratory findings must be applied.

Finally, reading researchers should consider strengthening their research programs by adding practitioners to their research teams. This does not mean the addition of *former* practitioners who are currently graduate students training to become researchers. Such people, of course, are expected to emulate researchers and, as such, can hardly be expected to represent their "old lives" as practitioners. Instead, practitioners should be hired *as practitioners* and should be charged with the responsibility of insuring relevance and reality in research designs and in the conclusions drawn from findings. This is not a "pie-in-the-sky" notion. Such "teacher collaborators" are an integral part of the research projects conducted at the Institute for Research on Teaching and their impact on the quality of our research has been significant.

Future Cooperative Ventures

Research on reading is a dynamic field; so is research on teaching. The nature of schooling dictates that the ultimate agenda for both communities is a shared one—applying findings in classrooms to improve instruction. This shared agenda could be expedited by collaborative ventures. While the possibilities for such cooperation are extensive, the following five examples are illustrative.

One of the differences heretofore separating the two research communities has been the respective object of focus. Reading researchers have focused on the curriculum and the learner; teaching researchers have focused on the teacher and instruction. Recently, however, each research community has become more interested in the other's object of focus as findings have pointed more and more to the complex interaction of learner variables, teacher variables and, in reading, text variables. As Brophy (1980b) points out, ". . . the development and merging of these interests should fuel a major trend in the 1980's." Certainly, reading and teaching researchers should collaborate in this trend.

Second, the affective goals of schooling will be a major research thrust in the next few years. Little is currently known about what constitutes effective teaching in this area. However, reading researchers, with their interest in comprehension as a personalized process and in affective responses to leisure reading, have much to contribute to collaborative research in this area.

Third, teacher effectiveness research is always bound by outcome measures. This is especially so in reading, as evidenced by past controversies regarding what constitutes achievement (see House *et. al.*, 1978 for one example) and by current debates regarding the validity of using standardized test measures in teacher effectiveness research. While the nature of the context will always exert some influence on the type of outcome measure used in future studies, additional effort will nevertheless be needed to identify appropriate outcome measures in reading, particularly in terms of the affective and personalized dimensions of language learning. This is another area which clearly calls for collaboration among reading and teaching researchers.

A fourth potentially profitable area of collaboration springs from the emerging findings on text processing and meta comprehension. Both imply the use of hueristics to help learners become more aware of and systematic about their language processing. Researchers from the reading and teaching communities could collaborate in this area to determine what role the teacher should play in presenting such hueristics to children. It would seem particularly important to insure that this not become yet another example of recommending patterned, technical, monitoring behavior to teachers.

Finally, the most important collaborative reading-teaching effort involves moving beyond structured opportunity to learn toward Carroll's concept of "quality of instruction." More specifically, the reading and teaching communities should conceptualize and test various guidance and assistance strategies which teachers of reading comprehension can use to go beyond the assessment tactics noted by Durkin (1979) and the recitation patterns noted by Duffy and McIntyre (1980), and toward what Barr and Dreeban (1977) describe as ". . . presenting subject matter clearly, visibly, audibly, and understandably . . . " and what Durkin (1979) defines as instruction. Although the reading profession has traditionally given little attention to this aspect of teaching, now is the time to do so.

The following five embryonic descriptions of instructional guidance may serve as a starting point. They represent a conceptual cycle. At the beginning point are those conceptions of teaching which involve little guidance beyond mechanical devices such as pacing and content covered (Barr, 1973-74). Here, the teacher is little more than a reactive responder who coordinates instruction by adjusting pace to learning rates of pupils. In addition to the fact that pacing is probably more a function of activity flow and management than of teaching (and, therefore, emphasizes the teacher's role as a technician), such conceptions of teaching offer little guidance to pupils. Providing little additional guidance but reflecting a less technical role for the teacher is the position espoused by Goodman (1979) and Smith (1975) who conceive of the teacher as a facilitator who avoids formal instruction and allows children to learn to read by interacting with books and language in natural, unconstrained and uncontrived ways. At the middle points of the cycle are two positions. Closest to the Goodman-Smith concept are Tierney and Pearson (1980) who emphasize the personalized nature of comprehension but also see the teacher as providing guidance in the manner of a ". . . tour guide on a journey to the center of the mind." Representing more teacher direction and assistance is the position described by Duffy and Sherman (1977) and by Duffy and Roehler (1980) in which the teacher is proactive in presenting cues and gradually diminishing assistance to guide pupils toward desired comprehension learnings. Finally, there is the conception of Becker, Engleman, and Carnine (1979) in which the teacher exercises extensive control in directing pupil responses, even to the extent of requiring unison responses from pupils. Hence, while this latter view may represent the ultimate in guidance, it also completes the cycle by bringing us back to the teacher as a technician who, in this case, often follows a prepared script when teaching.

While contextual influences suggest that no particular concept of teaching will be equally effective in all situations, it is intuitively obvious that teachers who provide assistance will create more success than those who simply structure "opportunity to learn" or who leave pupils to learn by natural means. Pursuit of such "quality instruction" is, in my view, the most important way in which the two research communities can collaborate in the future

CONCLUSION

Regardless of our particular backgrounds, reading and teaching researchers share a desire to create more effective schooling, particularly as it relates to language learning. Collaboration among the two research groups can help us reach this goal. Despite apparent disparities between reading educators' assumptions about classrooms and the findings from research on teaching, a "schema-match" based upon common "points of contact" is possible. Hopefully, this paper provides a beginning.

NOTES

1. It sometimes sounds like we are recommending that once one or another aspect of the reading process gets into the curriculum, the rest will take care of itself. Consider, for instance, Coltheart's (1976) dismissal of teachers in favor of curriculum when discussing early readers in a recent publication entitled "Reading Research: Advances in Theory and Practice":

   > . . . the fact is that reading teacher's practices are what they are. All readers, early or non-early, pass through a school curriculum, and if this curriculum is such that it eliminates the advantages that early readers initially had, that is that: early reading is simply beneficial.

ACKNOWLEDGEMENTS

The work reported herein is sponsored by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Teaching Division of the National Institute of Education, United States Department of Health, Education and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

The author gratefully acknowledges the contributions of Howard Seiler for searching the literature; of Linda Anderson, Jere Brophy, Joyce Putnam, Laura Roehler, George Sherman and Lee Shulman for critiquing earlier drafts of the paper; and of Peggy Medler and Theresa Wilbert for their care and patience in preparing the manuscript.

REFERENCES

Anderson, L., Evertson, C., & Brophy, J. An experimental study of effective teaching in first grade reading groups. *Elementary School Journal*, 1979, *79*, 193-222.

Anderson, L., Evertson, C., & Emmer, E. Dimensions in classroom management derived from recent research. *Journal of Curriculum Studies*, in press.

Artley, A. The teacher variable in the teaching of reading. *The Reading Teacher*, 1969, *23*, 239-248.

Au, K. H. *The comprehension-oriented reading lesson: Relationships to proximal indices of achievement*. A paper presented at the American Education Research Association, Boston, April, 1980.

Barnes, H., Putnam, J., & Wanous, D. *Multiple delivery systems*. Unpublished paper, Michigan State University/Lansing School District Teacher Corps Project, 1979.

Brophy, J. *Advances in teacher effectiveness research.* Occasional paper No. 18. Institute for Research on Teaching, Michigan State University, May, 1979.

Brophy, J. Advances in teacher research. *Journal of Classroom Interaction,* 1979, *15,* 1-7.

Brophy, J. *Recent research in teaching.* Invited address presented at the annual meeting of the Northeastern Educational Research Association, October, 1980.

Brophy, J. *Teacher behavior and its effects.* Occasional paper No. 25. Institute for Research on Teaching, Michigan State University, September, 1979.

Brophy, J. Teacher behavior and student learning. *Educational Leadership,* 1979, *37,* 33-38.

Brophy, J. *Teacher praise: A functional analysis.* Occasional paper No. 28. Institute for Research on Teaching, Michigan State University, October, 1979.

Brophy, J. *Teacher's cognitive activities and overt behaviors.* A paper presented at the conference of Basic Components in the Education of Mathematics Teachers, Rieste, West Germany, April, 1980.

Brophy, J., & Evertson, C. *Learning from teaching: A developmental perspective.* Boston: Allyn and Bacon, 1976.

Brophy, J., & Putnam, J. Classroom management in the elementary grades. In D. Duke (Ed.), *Classroom management.* Seventy-eighth yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1979.

Buike, S. *Teacher decision-making in reading instruction.* Research series No. 79. Institute for Research on Teaching, Michigan State University, May, 1980.

Buike, S., Burke, E. & Duffy, G. *Teacher conceptions of reading as they influence instructional decisions and pupil outcomes.* A paper presented at the annual conference of the International Reading Association, St. Louis, May, 1980.

Carroll, J. A model for school learning. *Teacher's College Record,* 1963, *64,* 723-733.

Cazden, C. Teaching as a linguistic process. In *Report on the task force on learning and teaching.* Washington National Institute of Education, 1977.

Clark, C., & Yinger, R. *The hidden world of teaching: Implications of research on teacher planning.* Research series No. 77. Institute for Research on Teaching, Michigan State University, 1980.

Doyle, W. Paradigms for research in teacher effectiveness. In L. S. Shulman, (ed.), *Review of research in education* (1) Vol. 5. Itasca, Illinois: Peacock, 1977.

Duffy, G. *The nature of instruction in primary grade reading.* A paper presented at the National Reading Conference, San Antonio, November, 1979.

Duffy, G., & McIntyre, L. *A qualitative analysis of how various primary grade teachers employ the structured learning component of the direct instruction model when teaching reading.* Research series No. 80. Institute for Research on Teaching, Michigan State University, June, 1980.

Duffy, G., & Roehler, L. *Classroom teaching is more than opportunity to learn.* Unpublished paper, Michigan State University, 1980.

Duffy, G., & Sherman, G. *Systematic reading instruction* (1) edition. New York: Harper and Row, 1977.

Duffy, G., Sherman, G., & Roehler, L. *How to teach reading systematically,* (1) edition. New York: Harper and Row, 1977.

Dunkin, J., & Biddle, B. *The study of teaching.* New York: Holt, Rinehart and Winston Inc., 1974.

Durkin, D. What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly,* 1979, *14,* 481-533.

Edmonds, R. Effective schools for the urban poor. *Educational Leadership,* 1979, 15-24.

Fenstermacher, G. On learning to teach effectively from research on teacher effectiveness. In C. Denham & A. Lieberman (Eds.), *Time to learn.* Washington: National Institute of Education, May, 1980.

Fenstermacher, G. *The value of education on teaching for teaching skill and teaching manner.* A paper presented at the annual convention of the American Educational Research Association, Boston, April, 1980.

Fisher, C. *Academic learning time as instructional feedback.* A paper presented at the annual convention of the American Educational Research Association, Boston, April, 1980.

Fisher, C., & Berliner, D. *Quasi-clinical inquiry in research on classroom teaching and learning.* Beginning Teacher Evaluation Study Technical Report. San Francisco: Far West Laboratory for Educational Research and Development, 1977.

Fisher, C., Berliner, D., Fliby, N., Marliave, R., Cahen, L., & Dishaw, M. Teaching behaviors, academic learning time and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), *Time to learn.* Washington: National Institute of Education, May, 1980.

Flanders, N. A. *Analyzing teacher behavior.* Reading, Ma.: Addison-Wesley, 1970.

Gage, N. L. *The scientific basis of the art of teaching.* New York: Teachers College Press, Columbia University, 1978.

Good, T. *Research on teaching*. An invited address presented at the annual convention of the American Educational Research Association, Boston, April, 1980.

Good, T., & Brophy, J. *Looking in classrooms*. New York: Harper and Row, 1978.

Good., T., & Grouws, D. The Missouri mathematics effectiveness project: An experimental study in fourth grade classrooms. *Journal of Educational Psychology*, (1979), *71*, 355-362.

Goodman, K. Comments made at the Conference on Basic Skills. National Institute of Education, Washington, D. C., September, 28, 1979.

Green, J. C. *A study of the effect on performance of cueing first grade children to the difficulty of a paired associate learning task*. Unpublished doctoral dissertation, Michigan State University, 1978.

Harste, J., & Burke, C. A new hypothesis for reading teacher research: Both the teaching and learning of reading are theoretically based. In P. D. Peterson (Ed.), *Reading: Theory, research, and practice*. 26th yearbook of the National Reading Conference. Clemson, South Carolina: National Reading Conference, 1977.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. F. No simple answer: Critique of the follow through evaluation. *Harvard Educational Review*, 1978, *48*, 128-160.

Jackson, P. *Life in classrooms*. New York: Holt, Rinehart and Winston, 1968.

Joyce, B. *Toward a theory of information processing in teaching*. Research series No. 76. Institute for Research on Teaching, Michigan State University, March, 1980.

Kamil, M., & Pearson, P. D. Theory and practice in teaching reading. *New York University Education Quarterly*, 1979, Winter, 10-16.

Koehler, V. Classroom process research: Present and future. *Journal of Classroom Interaction*, 1977, *13*, 3-11.

Kounin, J. *Discipline and group management in classrooms*. New York: Holt, Rinehart and Winston, 1970.

Lanier, J., & Shulman, L. *Technical proposal*. Institute for Research on Teaching, Michigan State University, 1976.

Leinhardt, G., Ligmond, N., & Cooley, W. *Reading instruction and its effects*. Paper presented at the annual convention of the American Educational Research Association, Boston, April, 1980.

McDonald, F. J. Research and development strategies for improving teacher education. *Journal of Teacher Education*, 1977, *28*, 29-33.

McDonald, F. J., & Elias, P. *The effects of teaching performance on pupil learning*. Final report, Volume I, Beginning Teacher Evaluation Study. Princeton, N.J.: Educational Testing Service, 1976.

Medley, D. Research on teacher effectiveness—where it is and how it got there. *Journal of Classroom Interaction*, 1977, *13*, 16-21.

Medley, D. *Teacher competence and teacher effectiveness: A review of process-product research*. Washington, D.C.: American Association of Colleges of Teacher Education, 1977.

Medley, D. The effectiveness of teachers. In P. Peterson and H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications*. Berkeley: McCutchan Publishing, 1979.

Medley, D., & Mitzel, H. Some behavioral correlates of teacher effectiveness. *Journal of Educational Psychology*, 1959, *50*, 239-246.

Morine-Dershimer, G. *Teacher plans and classroom reality: The South Bay Study, Part IV*. Research series No. 60. Institute for Research on Teaching, Michigan State University, 1979.

Osborn, J. *Do teachers really use the manual?* A paper presented at a conference on reading comprehension instruction, Center for the Study of Reading, University of Illinois—Urbana, May, 1980.

Peterson, R. Direct Instruction reconsidered. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications*. Berkeley: McCutchan Publishers, 1979.

Program on Teaching Effectiveness. An experiment on teacher effectiveness and parent-assisted instruction on the third grade. A symposium chaired by N. Gage at the annual convention of the American Educational Research Association, 1978.

Ramsey, W. An evaluation of three methods of teaching reading. In A. Figurel (Ed.), *Challenge and experiment in reading*. International Reading Association Conference Proceedings, 1962.

Richey, R. Durkin & Tierney keynote research conference. *Michigan secondary reading interest council newsletter*, 1980, *2*, 4-6.

Roehler, L., & Schmidt, W. Language arts. In *Progress report, January 1—June 30, 1980*. Institute for Research on Teaching, Michigan State University, 1980.

Rosenshine, B. V. Classroom instruction. In N. L. Gage (Ed.), *The psychology of teaching methods. Seventy-fifth yearbook of the National Society for the Study of Education, Part I. Chicago: Press, 1976*.

Rosenshine, B. Content, time and direct instruction. In H. Walberg & P. Peterson, (Eds.), Research on teaching: Concepts, findings and implications. Berkeley: McCutchan Publishing Company, 1979.

Rosenshine, R. How time is spent in elementary classrooms. In C. Denham & A. Lieberman (Eds.), *Time to learn.* Washington, D.C. National Institute of Education, May, 1980.

Rosenshine, B. *Teaching behaviors and student achievement.* London: National Foundation for Educational Research in England and Wales, 1971.

Rosenshine, B., & Berliner, D. Academic engaged time. *British Journal of Teacher Education,* 1978, *4,* 3-16.

Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R. Travers (Ed.), *Second handbook of research on teaching.* Chicago: Rand McNally, 1973.

Rupley, W. *Credible variables which relate to teacher effectiveness in reading instruction.* Paper presented at the annual convention of the National Council of Teachers of English, San Diego, November, 1975.

Rutter, M., Maughan, B., Martimore, P., Ouston, J., & Smith, A. *Fifteen thousand hours: Secondary schools and their effects on children.* Cambridge: Harvard University Press, 1979.

Schubert, W. Recalibrating educational research: Toward a focus on practice. *Educational Researcher,* 1980, *9,* 17-24.

Schwille, J., Porter, A., Gant, M., Belli, G., Floden, R., Freeman, D., Knappen, L., Kuhs, T., & Schmidt, W. *Factors influencing teachers decisions about what to teach: Sociological perspectives.* Research series No. 62. Institute for Research on Teaching, Michigan State University, October, 1979.

Shavelson, R. *Research on teacher's pedagogical thoughts, judgements, decisions and behaviors.* Unpublished paper, University of California at Los Angeles, 1980.

Shulman, L. *Reflections on individual differences and the study of teaching.* Unpublished paper, Michigan State University, 1980.

Sloat, K. *Characteristics of effective instruction.* Paper presented at the annual meeting of the American Educational Research Association, Boston, April, 1980.

Smith, F. *Comprehension and Learning: A conceptual framework for teachers.* New York: Holt, Rinehart and Winston, 1975.

Stallings, J. *Teaching basic reading skills in the secondary schools.* A paper presented at the annual convention of the American Educational Research Association, 1978.

Tharp, R. *The direct instruction of comprehension: Results and description of the Kamehameha Early Education Program.* Paper presented at the annual convention of the American Educational Research Association, Boston, April, 1980.

Tierney, R. and LaZansky, J. *The rights and responsibilities of readers and writers: A contractual agreement.* Reading education report No. 15. Center for the Study of Reading, University of Illinois-Urbana, January, 1980.

Tierney, R., Pearson, P. D. *New priorities for teaching reading in the eighties.* A paper distributed at the International Reading Association Pre-convention Institute on Comprehension Instruction, St. Louis, May, 1980.

Vinsonhaler, J. *The consistency of reading diagnosis.* Research series No. 28. Institute for Research on Teaching, Michigan State University, 1979.

Zigmond, N., Leinhardt, G., & Cooley, W. *Improving instructional environments for learning disabled students.* A paper presented at the annual convention of the American Educational Research Association, Boston, April, 1980.

LINDA M. MCHUGH
RAY R. BUSS
Cañada College

DIAGNOSTIC AND PRESCRIPTIVE DECISIONS MADE BY SPECIALISTS AND CLASSROOM TEACHERS: QUADRATIC ASSIGNMENT AS A RELIABILITY MEASURE

Only recently have researchers begun to study the diagnostic and prescriptive decisions made by specialists. Researchers have become interested in determining whether the faith that parents, administrators and classroom teachers are encouraged to have in the evaluations of children is justified, and in determining whether or not academic programs for diagnosticians need to focus more closely upon training professionals to make more reliable, or consistent, decisions. These are valid concerns since the ultimate purpose of a diagnostic and prescriptive evaluation is the implementation of a program of learning for a child who needs help. Aspects of such programs, what is or is not done for any particular child at school, are often the direct result of decisions made by a diagnostician or group of diagnosticians. The implications of such decision-making power over the present and future academic career of a child warrant study.

The purpose of the present study was to compare the diagnostic and prescriptive decisions made by reading specialists, learning disabilities teachers, Title 1 teachers and fourth grade classroom teachers to determine whether of not they identify similar reading skills weakness(es) and instructional needs of children, and prescribe similar reading programs for children.

Most of the previous research relating to the present study has been done at the Institute for Research on Teaching (IRT) on the Michigan State University campus. The overall results of the IRT studies and a master's thesis, one of the first studies of diagnostic and prescriptive decision making, suggest that there is little consistency among the diagnostic and prescriptive decisions made by diagnosticians.

An early IRT study investigated the problem-solving behaviors of select diagnosticians (Weinshank, 1978). Eight reading clinicians were asked to prepare diagnostic and remedial reports for three case studies. Two of the case studies were on different children; the third was a slightly altered version of one of the two studies. Weinshank found that over all case studies, diagnosticians were more likely to agree with the diagnostic and prescriptive decisions made by other diagnosticians than to agree with their own decisions about a single case study on separate occasions.

A second study considered the diagnostic and prescriptive agreement between different types of clinicians: reading specialists and learning disabilities teachers (Van Roekel, 1979). Van Roekel hypothesized that the specialist types were trained to pay attention to different diagnostic information about children when evaluating their skills and that these different perspectives would influence the diagnostic conclusions drawn by each type of specialist. In order to test this hypothesis, twenty subjects were asked to write a diagnostic summary for each of two case studies. The data from these summaries indicated that the subjects indeed tended to pay attention to different diagnostic information about each child. However, the different perspectives taken by the diagnosticians were not functions of their jobs descriptions.

A final IRT study examined the diagnostic decisions made by ten classroom teachers about children with reading problems (Gil, 1979). Gil found that there was little agreement among the diagnostic decisions made by classroom teachers. Indeed, classroom teachers agreed with each other less often than reading specialists. Furthermore, there was not strong agreement (.52) among reading specialists (Weinshank, 1978).

Outside of the IRT studies, one other study exists which examined the diagnostic and prescriptive decisions made by reading specialists (McBride, 1978). The purpose of this study was to examine the "usefulness of a diagnostic oral reading test as a prescriptive tool" as measured by the similarity among the prescriptive programs specified by graduate students in reading. McBride had twenty-three graduate students enrolled in a reading methods course make prescriptive decisions for each of four case studies. The case studies included information from the administration of the *Diagnostic Reading Scales* (Spache, 1972) to each case study child. McBride found that the corrective programs specified by her subjects were disparate. Therefore, she concluded "that the diagnostic information given to the subjects did not suggest the same course of instruction to all subjects" (p. 29).

McBride's conclusion seems to be substantiated by the research done at IRT. However, all of these studies exhibit similar limitations: biased sampling and limited statistical tests. Each IRT study was based upon a very small sample of the population of interest. In the case of specialists, the sample was select and not representative of the population of specialists who work in public school systems. The subjects were all recognized as experienced, at least five years work as a clinician, and considered exceptionally reputable as clinicians. Many of the IRT subjects were reading faculty members at Michigan State University. Most had or were obtaining doctoral degrees. McBride's sample was likely to be more homogeneous than the population of reading specialists working in public schools because it was comprised of graduate students enrolled in the same reading methods course at the same university.

The conclusions drawn from the IRT studies were based upon correlation coefficients. The significance of the correlation is a matter of subjective interpretation. Furthermore, correlational measures only relate two aspects of decision making at any given time. To more accurately consider the reliability of the decisions diagnosticians make about a child, it is essential to examine the profile of the decision rather than individual or pairwise components of the decision process. The quadratic assignment procedure, a non-parametric statistical test, allows the researchers to examine the similarity of decisions made by two groups. Thus, quadratic assignment was used to analyze the data in the present study.

METHOD

A sample of ninety subjects was randomly selected from each of four populations: reading specialists, learning disabilities teachers, Title 1 teachers, and fourth grade classroom teachers in the state of Wisconsin. These populations were included in the study because they include personnel who typically make diagnostic and prescriptive decisions about children with reading problems in the schools. Sampling frames were obtained from the Wisconsin Department of Public Instruction. Materials for the study were mailed to the subjects, they were asked to complete the qustionnaires and return them by mail. The return rate was about 46% over each population for each part of the study.

Materials

The present study included two parts. In Part A, subjects were asked to make diagnostic and prescriptive decisions about the child profiled in Case Study A. This child had a predominant reading skill weakness in comprehension. In Part B, the same subjects were asked to make diagnostic and prescriptive decisions about the child represented in Case Study B. This child had a predominant reading skill weakness in word identification.

Case Studies

One case study was developed for each of the two parts of this research. Case study reports allow replicability, constraint on extraneous variables, and inclusion of a greater number of subjects to enhance generalizability. For these reasons, the use of case study reports seemed the most reasonable way of having the subjects receive information and make diagnostic and prescriptive decisions about children experiencing problems learning to read.

Each case study was based upon the actual assessment of a fourth grade child with reading problems. The first section of the case study report included the child's age, grade level, placement in school and scores on formal and informal tests. Test scores included the results of an informal reading inventory (IRI), a test of phonic skills, the *Johnson Basic Sight Vocabulary Test* (Johnson, 1976), *The Woodcock Reading Mastery Tests* (Woodcock, 1973), and the *Diagnostic Reading Scales* (Spache, 1972). These measures were selected because they often are used for diagnosis in Wisconsin public schools.

The second section of each case study contained observational data. The purpose of observational data was to report behaviors of the child which may expand upon or help explain the child's test scores. Observational information may also suggest overused or inefficient strategies the child may use when reading which directly relate to the reading problem.

The final section of each case study, "additional information", included information such as school, health, and developmental history. Again, such information may help the diagnostician understand the child's reading problem more fully.

DIAGNOSTIC RESPONSE QUESTIONNAIRE

The Diagnostic Response Questionnaire consisted of twelve major questions covering twenty-six diagnostic and prescriptive decisions to be made about the case study child, the job description of the subject, and demographic information concerning the experience and educational background of the subject. Each question about diagnosis and prescription was broken down into an exhaustive list of responses. The subject was asked to decide "yes" or "no" for each question. Each response was used as a component of the subject's profile of responses. This means that twenty-six specific decisions, rather than twelve general ones, were analyzed.

Data Analysis

Data from each questionnaire were organized into matrices, called proximity matrices, for each job description group. The proximity matrices were then compared in a pairwise fashion using the Q/AP computer program (Baker, Hubert, and Schultz, 1977). The procedure allows the researcher to get more than a "feel" for the qualitative data. It allows the reliability of the decisions made between diagnostician groups to be statistically tested. Moreover, the entire profile of a diagnosis and prescription is tested rather than discrete parts of the evaluation. Briefly, Q/AP allows an observed value, Γ observed, to be compared to a reference distribution of Γ' values, generated by the successive permutation of the matrices being compared. One matrix, arbitrarily named the structure matrix, is held constant while the rows and columns of a second matrix, known as the proximity matrix, are permuted. If the matrices are very similar, i.e., if high frequencies are in corresponding cells in the two matrices, then the largest value will be Γ observed. If, on the other hand, the matrices are dissimilar, then Γ observed will be small relative to the distribution of the generated Γ' values.

A numerical example of the Q/AP procedure has been provided below (see Table 1). A portion of the matrices of responses from reading specialists and Title 1 teachers for Case Study A will provide the example data. The numerals in each cell reflect how many subjects in the diagnostician group said "yes" to both items on the qustionnaire. Notice that zero is entered on the diagonals and that the matrices are symmetrical.

TABLE 1

	5a	5b	5c	5d	5e
5a	0	4	8	14	15
5b	4	0	2	3	3
5c	8	2	0	14	12
5d	14	3	14	0	29
5e	15	3	12	29	0

Reading Specialists

	5a	5b	5c	5d	5e
5a	0	1	15	22	18
5b	1	0	1	1	1
5c	15	1	0	15	12
5d	22	1	15	0	24
5e	18	1	12	24	0

Title I Teachers

The first step in the Q/AP process is to multiply the corresponding cells in each matrix, and then sum these crossproducts to obtain Γ observed. The observed Γ for the two matrices above is 3520. The next step is the generation of the distribution of Γ' values against which Γ observed will be compared for significance. As stated earlier, this distribution is obtained by holding the first matrix constant and interchanging one set of rows and the corresponding set of columns in the second matrix, crossmultiplying on a cell by cell basis, and summing the crossproducts to obtain Γ'. This procedure is repeated many times to obtain a whole distribution of Γ values (Baker, Hubert, and Schultz, 1977). In the present study, 750 permutations were used to obtain the reference distribution. One sample of this part of the process is given below (see Table 2). Here rows 5b and 5c of the reading specialist matrix were interchanged and simultaneously columns 5b and 5c were permuted while the Title 1 matrix was held constant. The crossproducts were calculated and Γ' was obtained. Γ' was 2902 for this permutation.

TABLE 2

	5a	5c	5b	5d	5e
5a	0	8	4	14	15
5c	8	0	2	14	12
5b	4	2	0	3	3
5d	14	14	3	0	29
5e	15	12	3	29	0

Reading Specialists

	5a	5b	5c	5d	5e
5a	0	1	15	22	18
5b	1	0	1	1	1
5c	15	1	0	15	12
5d	22	1	15	0	24
5e	18	1	12	24	0

Title I Teachers

TABLE 3
Results of Pairwise Comparisons of Responses by Diagnostician Groups for Case Study A

Proximity Matrix	vs.	Structure Matrix	Probability
Reading Specialists	vs.	L. D. Teachers	$p < .00133$
Reading Specialists	vs.	Title I Teachers	$p < .00133$
Reading Specialists	vs.	Classroom Teachers	$p < .00133$
Title I Teachers	vs.	L. D. Teachers	$p < .00133$
Title I Teachers	vs.	Classroom Teachers	$p < .00133$
L. D. Teachers	vs.	Classroom Teachers	$p < .00133$

The final step in using the quadratic assignment technique is to decide whether or not the Γ observed value occurs far enough in the tail of the generated distribution to allow us to conclude that the matrices are similar. In the present study, an overall critical value of $p<.05$ was established. This means that the necessary critical value for each of six pairwise comparisons was $p>.00833 = .05/6$. In each case (see Tables 1 and 2), the Γ observed exceeded the largest Γ' value in the reference distribution. It was determined that $p<.00133$ for each comparison. Consequently, it was concluded that the observed similarity between pairs of diagnostician groups was very unlikely to be due to chance.

RESULTS

The results of each pairwise comparison are displayed in Table 3 for Case Study A and Table 4 for Case Study B. In each case, the *p* values, the likelihood that the observed relationship was the result of chance response, occurred at

TABLE 4

Results of Pairwise Comparisons of Responses
by Diagnostician Groups
for Case Study A

Proximity Matrix	vs.	Structure Matrix	Probability
Reading Specialists	vs.	L. D. Teachers	$p < .00133$
Reading Specialists	vs.	Title I Teachers	$p < .00133$
Reading Specialists	vs.	Classroom Teachers	$p < .00133$
Title I Teachers	vs.	L. D. Teachers	$p < .00133$
Title I Teachers	vs.	Classroom Teachers	$p < .00133$
L. D. Teachers	vs.	Classroom Teachers	$p < .00133$

$p<.00133$, which allows an overall error rate of .05. In other words, the measure of similarity for each pair of diagnostician groups indicated that the groups made very similar decisions. Therefore, the reading specialists, learning disabilities teachers, Title 1 teachers, and classroom teachers, made similar decisions about the general reading skill weakness, remedial needs and instructional approaches to remediation to be used with the children profiled in each case study report.

CONCLUSION

The results of the present study are particularly important because they contradict the conclusions drawn by previous researchers about the reliability of the diagnostic and prescriptive decisions made by specialists and teachers. Another important aspect of the present study is the contribution of a test which allows researchers to compare the decisions of diagnosticians over many qualitative components and to test these decisions for statistical significance.

Q/AP represents an important contribution to the array of choices for the analysis of qualitative data. It allows the researcher to actually test an hypothesis and to treat multiple measure qualitative data in much the same way quantitative data are treated by inferential methods. In the present study, Q/AP has allowed us to conclude with confidence that specialists and teachers sampled for this study made very similar diagnostic and prescriptive decisions.

REFERENCES

Baker, F. B., Hubert, L. J., & Schultz, J. *Quadratic assignment program Q/AP*. (Laboratory of Experimental Design). Unpublished manuscript, University of Wisconsin-Madison, 1977.

Gil, D. *The clinical problem solving behavior of classroom teachers as they diagnose a child's* *reading performance in experimental and classroom situations*. Unpublished doctoral dissertation, Michigan State University, 1979.

Johnson, D. D. *The Johnson basic sight vocabulary test*. Lexington, Mass.: Ginn and Company, 1976.

McBride, E. *An investigation of the reading program prescriptions made by different teachers given diagnostic information obtained from the diagnostic reading scales*. Unpublished master's thesis, University of Wisconsin-Madison, 1976.

Patriarca, L., Van Roekel, J., & Lezotte, L. *Simulated reading and learning disability cases: effective tools for research and teacher education* (Tech. Rep. 29). East Lansing: Michigan State University, The Institute for Research on Teaching, 1978.

Spache, G. D. *Diagnostic reading scales* (rev. ed.). Monterey: CTB/McGraw-Hill, 1972.

Van Roekel, J. R. *The problem-solving behavior of teachers trained in reading and teachers trained in learning disabilities when diagnosing simulated cases in educational difficulty in a clinical setting*. Unpublished doctoral dissertation, Michigan State University, 1979.

Weinshank, A. B. *The relationship between diagnosis and remediation in reading: a pilot study* (Tech. Rep. 37). East Lansing: Michigan State University, The Institute for Research on Teaching, 1978.

Woodcock, R. W. *Woodcock reading mastery tests, form A*. Circle Pines, Minn.: American Guidance Service, Inc., 1973.

SANDRA STOTSKY
Curry College

TOWARD REASSESSMENT OF THE PRINCIPLES UNDERLYING THE CHOICE OF VOCABULARY AND THE TEACHING OF WORD ANALYSIS SKILLS IN READING INSTRUCTIONAL MATERIAL

INTRODUCTION

The importance of vocabulary in reading comprehension has been consistently reported in several areas of reading research. Chall (1958), in a critical review of readability formulas, found that a measure of vocabulary load was the major factor in almost all readability formulas. Studies of children's language development show a high correlation between pre-schoolers' knowledge of word meanings and achievement in reading at higher grade levels (e.g., Loban, 1970). Factor analyses of component skills in reading comprehension point to a knowledge of words as the essential component in reading comprehension (e.g., Davis, 1971). Further, a summary of research findings from studies in the teaching of vocabulary (Petty, Herold, & Stoll, 1968) indicates that some systematic attention to vocabulary teaching is better than no attention at all.

An examination of the way in which new words are introduced in reading series for the elementary grades (Harris & Jacobson, 1973-1974) indicates that vocabulary seems to be introduced primarily on the basis of frequency of use in written material. An inspection of several recent reading series (listed below) suggests that vocabulary is also determined by idiosyncratic usage in literary selections. When literary selections are adapted, there appears to be no indication of the principles followed in the choice of words. It is highly rational in beginning readers to teach children to read words that are in their oral vocabulary and that are among the most frequent words in written material as well. However, too much reliance on the principle of frequency beyond the decoding stages or on exposure to a richer (and possibly uncontrolled) variety of words in literacy selections may not be sound from a long-range point of view if it precludes the possibility for more systematic development of a reading vocabulary.

In 1974, O'Rourke (1974) proposed a planned program of vocabulary development that contains as its nucleus the study and use of generative roots and affixes. Such a program would foster the systematic expansion of children's knowledge of words by enabling students to transfer the meaning of an element learned in one word to other words containing that element. However, O'Rourke did not investigate whether or not current reading instructional material provides opportunities for expanding children's knowledge of words on this basis.

PURPOSE

This study sought to discover how comprehensively and systematically vocabulary is taught in reading instructional material. Because it was not possible to deal with all categories of word elements, the scope of this research was limited to the teaching of prefixes and the use of prefixed words; this category could be dealt with comprehensively. Six widely-used reading series and their accompanying workbooks (Grades 2 to 6) were surveyed to determine: (1) how prefixes are taught, (2) what prefixes are taught, (3) in what order, (4) at what grade levels, and (5) how many different exemplars of the prefix are offered in meaningful reading material. Findings from this survey could provide data with which to judge whether or not these series provide opportunities for systematic expansion of children's knowledge of words.

PROCEDURES

The first issue was to decide upon the definition of the term "prefix" to be used. According to Marchand's (1969) text, considered by Aronoff (1976) as the most comprehensive text on the subject of English word-formation, prefixes are defined as "bound morphemes which are preposed to free morphemes" (p. 129). Marchand emphasizes that "only such particles as are prefixed to full English words of general, learned, scientific or technical character can be termed prefixes" (p. 132). For example, *in-* as in *invisible* and *insane* is a prefix, but *in-* as in *intent* is not, as *-tent* is not an independent English word in the sense required. By definition, then, the elements *ad-* in *adjacent, ex-* in *expect, pre-* in *prefer, com-* in *companion, ob-* in *obstacle,* etc. are not prefixes in these words because they are not preposed to independent words. They are etymological elements attached to roots in non-composite words (i.e., words that cannot be analyzed on the basis of English word-formation). Indeed, such elements as *com-, ob-,* and *ad-* are never prefixes. Cronnell (1973) uses the term "pseudoaffixes" to describe these initial elements.

Second, the following six reading series were selected for the survey:

(1) Allyn & Bacon Basic Reading Series (1968)
(2) Macmillan Reading Program, Revised Edition (1970)
(3) Ginn 360 Reading Series (1970)
(4) Scott Foresman Reading Systems (1971-1972)
(5) Holt Basic Reading System (1973)
(6) Ginn 720 Reading Series (1976)

They were chosen because: (1) they have been among the most widely-used series in this country; (2) they range in terms of date of publication from the 1960's through the 1970's, thus reflecting differing theoretical issues over the past decade influencing the construction of reading instructional material; and (3) they represent a rough balance between those tending more to use adapted or constructed selections and those tending more to use unadapted literary selections. To judge from the information in the acknowledgment pages at the beginning of the readers, the Allyn & Bacon, Macmillan, and Ginn 360 series appear to contain a rough balance between adapted and unadapted selections; the Holt and Scott-Foresman series tend to contain more unadapted literary selections; the Ginn 720 series seems to fall between these two groups. Although there tends to be an increasing number of informational selections at higher grade levels in every series, most selections at each grade level are fictional or literary in nature.

Third, all the reading material in the pupil readers and workbooks at each grade level in all series was read word by word. For each grade level in each series, tables were constructed containing all exemplars of prefixes taught at that grade level and at previous grade levels. It should be noted that the listing of a word in the tables indicated only the appearance of a prefixed word (including any of its derivatives), not frequency, since the focus of the study was on opportunities for expansion of word knowledge through transfer of meaning to similarly prefixed words. The data in the tables were drawn only from the material in the readers and workbooks that required either meaningful reading or writing of prefixed words; this criterion excluded words listed for purposes of alphabetizing, stress placement, syllabication, etc.

FINDINGS

Table 1 indicates the total number of different exemplars for each prefix at each grade level once a prefix is introduced or mentioned in the teachers' guides, readers, or workbooks in the six series. The numeral *zero* indicates that no exemplars were found at that grade level even though the prefix was introduced or mentioned at that level or at an earlier one. Considerable variation in the order of introduction exists for most prefixes; only *un-*, *re-*, *dis-*, *in-*, and *im-* are introduced in all series by Grade 4. Many prefixes are not introduced at all. It is not clear from the reading series themselves or from any research literature what rationale might provide the basis for the choice and order of introduction of prefixes in general. It is clear only that many common prefixes used in elementary reading material are not taught in all series.

Table 1 also reveals a paucity or total absence of exemplars for many prefixes at many grade levels throughout these series. The following reasons are offered to account for this scarcity:

a. There seems to be a lack of coordination between what is suggested for teaching (or mentioned) in the teachers' guides and what is available in the corresponding reading selections in the readers or workbooks. Quite often prefixes are suggested for teaching in exercises in the guides but only one or two exemplars (if any at all) appear in the reading selections. In general, very few corresponding workbook pages contain an exercise for a prefix suggested for teaching in the guide.

b. There appears to be a misunderstanding of prefixation in all the reading series, according to the definition that was followed in this research (see Stotsky, 1976, for a grade-by-grade examination of the way in which each series defines, exemplifies, and teaches prefixes). Distinction is usually not made between prefixed words, such as *remake, precaution,* or *defrost,* and words with initial etymological elements, such as *reflect* or *deliver*. As a result, many words offered as examples of prefixed words are incorrect or misleading (e.g., *disappoint)* and therefore confusing, difficult, or useless for teaching or learning purposes.

A comparison of the total number of different exemplars for all prefixes at each grade level in each series reveals some differences between the series, but there appears to be no relationship between the total number of different exemplars and whether the series contain a balance between adapted and unadapted literary selections or mainly unadapted ones.

CONCLUSIONS AND IMPLICATIONS

Even though all the prefixes listed in Table 1 were not taught in all series, it seems reasonable to suggest that they should all be taught to all elementary

TABLE 1

Number of Different Exemplars of Each Prefix After the Prefix is First Introduced or Mentioned in Six-Reading Series

Series	**Allyn & Bacon**					**Macmillan**					**Ginn 360**					**Scott-Foresman**					**Holt**					**Ginn 720**				
Grade	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
anti-			1	0	2				3	4									3	0				1	1					1
circum-															1									1						
co-										2										6					3					
counter-			1	1	2																			1	1					
de-				4	0						2	1	0	3	2						0	1	0	2	4	0	3	0	4	4
dis-		6	7	14	11		7	8	18	27			9	15	21		3	7	11	14	1	2	7	7	23		9	4	11	19
em-					1					3																				
en-			1	11	6				12	15			7	13	11			2	12	12	2	1	9	9	15			4	16	12
extra-																									0					
fore-									12	5				13	4				9	8		2	5	6	5				13	9
il-										1														0	1					
im-		1	2	6	9		4	5	6	6			4	5	8			2	4	7			1	2	8			2	5	8
in-		3	1	8	8		0	4	16	20			5	10	17			3	5	18		1	6	3	14			[illegible]	13	13
ir-										3								0	2	2				2	0					
inter-				1	1					2										5					4					3
intra-																									2					0
mid-									11	7									3	3		2	3	3	3					
mis-			1	2	7				6	5								2	3	3		0	2	4	4					
non-			0	1	2					10			0	3	6				3	1		0	2	5	8			1	4	5
post-															0					0				0	0					0
pre-			0	1	3				6	3		1	1	3	5			1	2	4			3	3	7	1	4	2	2	7
pro-			0							0				0	0								0	0	0				0	0
re-		2	8	25	19		6	17	19	27	1	12	10	9	13		5	6	21	17		4	11	13	14	1	6	10	9	22
semi-					1																				2					
sub-			1	2	3				3	2			4	3	6					3	1	0	3	3	5				3	3
super-							4	0	0	5									2	2			4	7	3				1	5
trans-				3	0					2										2				3	2				1	4
un-		20	27	53	71	12	23	60	90	100		34	56	63	97	7	14	26	59	75	4	10	40	31	69		23	37	63	86
TOTAL		32	52	132	144	12	44	94	202	249	3	48	96	140	191	7	22	49	139	182	8	23	96	106	198	2	45	64	145	201

school students. A majority of them were found in many different words in the six series, and almost all of them are used with many common words in our language. While Groff (1972) concluded from a review of the literature that there is no evidence to support the teaching of prefixes as a way to improve children's reading vocabulary, research reported by Graves & Hammond (1980) provides evidence that teaching prefixes correctly to Grade 7 students enabled both higher and lower ability experimental students to unlock the meaning of novel words better than a control group of students being taught vocabulary by a whole word method.

From an inspection of the data summarized in Table 1, it is also possible to judge that sufficient opportunities for systematic vocabulary development with respect to many prefixes do not exist at many grade levels in these series, whether or not the series contain mainly unadapted literary selections or a balance between adapted and unadapted ones. At present, choice of vocabulary seems to be governed either by idiosyncratic usage in literary selectons or by the principle of frequency or the degree of regularity in sound to symbol correspondence. If one may generalize from the findings of this study, the analysis of these six reading series suggests the need to reassess the rationale for choice of vocabulary in reading instructional material. Such a reassessment logically involves the question of what kinds of reading selections should constitute reading instructional material, since the vocabulary used in a specific mode of written discourse is organically related to the nature of that mode.

It is possible that literary selections, whether adapted or not, cannot provide sufficient opportunities for systematic vocabulary development because of their very nature. Vocabulary can be taught systematically if key words or words associated with these key words are repeated and if derivatives can be used. Repetition of key words does not typically occur in fictional, literary prose, nor do derivatives play any important role in its vocabulary. On the other hand, repetition of important words and the frequent use of affixed words are dominant features of informational, non-literary prose (Stotsky, 1979; Stotsky; 1981; Stotsky, in press). It is possible that the vocabulary of informational prose may be even more important for children to learn than the vocabulary of literature. If this is so, then reading educators should consider the use of another rationale for choosing reading selections in reading instructional material than the one used to guide the choice of reading selections in most current reading series. Venezky (1981) recently estimated that about 95% of the reading selections in current reading series are narrative fictional in nature.

There are also several areas of investigation for researchers to pursue. First, how widespread is the misconception of prefixation found in these six series? A preliminary inspection of several other reading series, supplementary reading instructional materials, and language arts textbooks for both the elementary and secondary levels reveals the same inaccuracies. Second, how accurate is the explanation and description of word analysis in professional textbooks on the teaching of reading and language arts? An examination of several texts (e.g., Dallman, Rouch, Chang, and DeBoer, 1978, p. 184, and Lapp & Flood, 1978, p. 256) also reveals inaccuracies. Thus an assessment of the way the topic of word analysis is presented and discussed in informational material for researchers and teachers as well as in instructional material for students is strongly recommended.

REFERENCES

Aronoff, M. *Word formation in generative grammar*. Linguistic Inquiry Monographs. Cambridge: MIT Press, 1976.

Chall, J. *Readability: an appraisal of research and application*. Columbus, Ohio: Ohio State University Press, 1958.

Cronnell, B. *Reading words of more than one syllable*. (Professional Paper No. 24.) Los Alamitos, Cal.: SWRL Educational Research and Development, 1973.

Dallman, M., Rouch, R., Chang, L., & DeBoer, J. *The teaching of reading, fifth edition*. New York: Holt, Rinehart and Winston, Inc., 1978.

Davis, F. Psychometric research on comprehension in reading. In F. B. Davis (Ed.), *The literature of research in reading with emphasis on models*. East Brunswick, N.J.: Iris Corp., 1971.

Graves, M. & Hammond, H. A validated procedure for teaching prefixes and its effects on students' ability to assign meaning to novel words. In M. Kamil & A. Moe (Eds.) *Perspectives on reading research and instruction*. Twenty-Ninth Yearbook of the National Reading Conference. Washington, D.C.: National Reading Conference, Inc., 1980.

Groff, P. Teaching affixes for reading improvement. *Reading Improvement*. Spring 1972, 28-31.

Harris, A., & Jacobson, M. Some comparisons between the basic elementary reading vocabularies and other word lists. *Reading Research Quarterly*, 1973-1974, *9* (1), 87-109.

Lapp, D. & Flood, J. *Teaching reading to every child*. New York: Macmillan Publishing Company, Inc., 1978.

Loban, W. *Stages, velocity, and prediction of language development: Kindergarten through grade twelve*. (Final Report) Washington: Office of Education, Bureau of Research, U.S. Department of HEW, 1970.

Marchand, H. *The categories and types of present-day English word-formation*. (2nd revised edition) Munich, Germany: C. H. Beck, 1969.

O'Rourke, J. *Toward a science of vocabulary development*. The Hague: Mouton Press, 1974.

Petty, W., Herold, C., & Stoll, E. *The state of knowledge about the teaching of vocabulary*. Champaign, IL: National Council of Teachers of English, 1968.

Stotsky, S. *Towards more systematic development of children's reading vocabulary in developmental reading programs for the middle to upper elementary grades*. Unpublished doctoral dissertation, Harvard University, 1976.

Stotsky, S. Teaching the vocabulary of academic discourse. *Journal of Basic Writing*, 1979, *2* (3), 15-39.

Stotsky, S. *Lexical cohesion in expository essay writing*. Paper presented at the annual convention of the International Reading Association, New Orleans, Louisiana, April 1981.

Stotsky, S. The vocabulary of essay writing: Can it be taught? *College Composition and Communication*, in press.

Venezky, R. *The origins of the present-day chasm between adult competency and literacy instruction*. Paper presented at the Conference on Literacy and Competency, Harvard Graduate School of Education, May 1981.

READING SERIES

Centennial Edition of the Sheldon Basic Reading Series. W. D. Sheldon, et al. Allyn & Bacon, Inc., 1968.

Holt Basic Reading System. E. Evertts, L. C. Hunt, & B. Weiss. Holt, Rinehart & Winston, Inc., 1973.

Macmillan Reading Program, Revised Edition. A. J. Harris, et al. Macmillan Co., 1

Reading 360. T. Clymer, D. Gates, et al. Ginn & Co., 1970.

Reading 720. T. Clymer, et al. Ginn & Co., 1976.

Scott Foresman Reading Systems. I. E. Aaron, et al. Scott, Foresman & Co., 1971.

JAMES V. HOFFMAN
SHARON O'NEAL
University of Texas at Austin

CHRISTOPHER BAKER
Texas Wesleyan College

A COMPARISON OF INSERVICE AND PRESERVICE TEACHER VERBAL FEEDBACK TO STUDENT MISCUES ACROSS TWO DIFFICULTY LEVELS OF TEXT

Teacher guided oral reading persists as a common practice in most elementary classrooms. Howlett and Wientraub (1978) report that over four-fifths of primary grade teachers responding to their survey engage students daily in oral reading exercises. Other than the research evidence pointing to the existence and wide use of oral reading in classrooms, we have no substantial history of research into the characteristics or effects of oral reading instruction on pupil performance. While emotional indictments of teacher guided oral reading appear from time to time in the professional literature (e.g., Artley, 1972), the practice continues. Making the situation even more unsettling is some recent classroom research pointing to positive outcomes for students in classrooms where teachers engage students frequently in oral reading (Anderson & Evertson, 1978; Stallings, Needels, & Stayrook, 1979). The void in our understanding for the place of teacher guided oral reading in a developmental program is sorely in need of attention.

In an instructional context oral reading takes on the form of a dialogue in which information is exchanged between teacher and student. These verbal interactions typically arise from the teacher's efforts to give students feedback about their miscues. Clements and Hoffman (1981) have found that over 35% of teacher talk during guided oral reading is miscue focused. What kinds of feedback do teachers typically rely on in such contexts? How do situational or setting variables influence changes in the kinds of feedback offered? What effects (both short and long term) do these different strategies have on the development of pupil competencies? Only recently has the research literature come to offer any information useful in answering any of these questions (Niles, 1980). It was in an effort to contribute to our knowledge in this area that the present study was conceived.

This research represents an attempt to systematically study the characteristics of teacher feedback to student miscues during oral reading. The research builds directly on the work of such researchers as Brophy and Good (1969), Goodman (1973), Brady and Lynch (1976), and Allington (1978). It is based on the conceptual framework for studying teacher feedback proposed by Hoffman (1979) who identified three critical dimensions to be considered in examining verbal feedback during oral reading: (1) selectivity—which miscues (or what proportion) are responded to by the teacher; (2) timing—when, or at what point, is the feedback offered; and (3) form—what are the qualitative characteristics of the feedback itself.

METHOD

Subjects

The subjects for this study were teacher/pupil dyads. Thirty-four elementary pupils were selected at random from students enrolled in a summer reading program at The University of Texas at Austin. The actual grade placement levels of the students were distributed evenly among grades one through five. The teachers were eighteen experienced classroom teachers enrolled in a graduate reading methods class and sixteen undergraduate education majors (inexperienced teachers) enrolled in their first reading methods course. Pupils were randomly assigned to teachers to form instructional dyads. Teachers and pupils had no instructional contact or familiarity prior to participation in this study. This was done so as to control for the possibility that prior knowledge of student needs might influence response patterns by teachers.

Procedures

Each student's approximate instructional reading level (92-98% Word Accuracy) was determined during a screening phase using an informal reading inventory developed from passages found in each of the basal readers of *The New Basic Readers* (Scott, Foresman, and Company, 1964). Reading achievement levels, as reflected on the informal reading inventory (IRI), were generally distributed evenly above and below grade placement.

Selected portions of basal readers in the *New Basic Reading Series* that had not been included in the informal reading inventory, were used as reading materials in this study. Each pupil read aloud to a teacher for approximately ten minutes from one section of text at an instructional level and for an additional ten minutes from a second piece of text at the next higher level within the series. The difficulty sequence was counter-balanced between subjects (i.e., easy first/hard second and hard first/easy second). All sessions were videotaped from concealed audiovisual equipment. Prior to commencing the oral reading sessions, identical sets of directions were given. Experienced and preservice teachers were informed that the purpose of the study was to record and examine interactions between teachers and students during oral reading instruction and that they should feel free to assist the student in as natural a manner as possible. The pupils were told they would read two texts aloud with a teacher present to guide them.

Coding

Videotapes were coded using the FORMAS-dyadic taxonomy (Hoffman & Baker, 1980). Coders were trained to use FORMAS to classify audiovisual recordings of the student-teacher interactions during oral reading instruction. Four major clusters of teacher/pupil interactive behaviors were monitored and analyzed for this study. See Table 1. Interrater reliability during coding was monitored with random checks for coder agreement. The coefficient of interrater agreement for nominal scores K, was the measure used to estimate the proportion of joint judgements of reading miscues after chance agreement was excluded (Hoffman, Gardner & Clements, 1980). For each dyad the reliability coefficients for agreed miscues ranged between .83 and 1.00. Interrater reliability coefficients for each category of behavior across the agreed upon miscues ranged between .79 and .96. Only single word miscues were coded and analyzed

TABLE 1

Four major clusters of teacher/pupil interactive behaviors

I. Miscue
   - A. Type: insertions; omissions; hesitations; substitutions; mispronunciations; calls for help; and repetitions.
   - B. Meaning change: high and low.
   - C. Syntactic acceptability: high; some; and low.
   - D. Grapho-phonic similarity: high and low.

II. Reaction (student's immediate behavior following miscue)
   - A. Type: repeated attempt; continuation; immediate self-correction; pause; call for help; and no opportunity.

III. Teacher Verbal Feedback
   - A. Type: no verbal; terminal (giving the text word); and sustaining (helping student to identify text word).
   - B. Form of sustaining: attending (noncue focusing); simple graphophonic; simple context; complex graphophonic (i.e., graphophonic followed by context); and, complex context (i.e, context followed by graphophonic).
   - C. Timing of teacher feedback: immediate (0 to 3 seconds); delayed (more than 3 seconds).
   - D. Point of teacher feedback: before the next sentence break; at the next sentence break; or after the next sentence break.

IV. Resolution: teacher identified text word; student identified text word; or miscue left unidentified.

in this study. Multiple miscues (similar to Weber's (1970) "scrambles") involving two or more contiguous text words were simply tallied.

RESULTS AND DISCUSSION

A total of 1,837 miscue interactions were coded. The average accuracy of oral reading for students in the easy material was about 85% and in the difficult material about 81%. Although these error rates (15% and 19% respectively) are high in comparison to the criteria used with the screening IRI, two important differences must be kept in mind. First, in the experimental setting hesitation and repetitions were always counted as miscues. This was not always the case in the screening IRI. Second, miscues which were self corrected in the experimental setting were counted. They were not counted in the screening IRI. Multiple miscues accounted for about 9% of the total. These miscues were not included in the analyses to be reported.

Selection: Which miscues did teachers respond to?

Teachers made some form of overt verbal response to only 37% of the single word miscues made by students. This figure roughly replicates the findings of Allington (1978) in his study of classroom oral reading instruction. Further analyses of our data revealed that teachers were more likely ($p < .05$) to respond to miscues made in difficult rather than easy material (40% versus 34%). Inservice teachers were also more inclined ($p < .05$) to respond than preservice teachers (40% versus 34%).[1]

---

[1] All tests of statistical significance were made following the model proposed by Castellan (1965) for the use of chi-squares in the partitioning of contingency tables.

Teachers were more likely ($p < .05$) to respond to miscues which affected meaning substantially (44% were responded to) than those which resulted in minimal meaning change (only 19% were responded to). Teachers seemed also to be sensitive to the ways in which students were reacting to their own miscues in determining whether or not to respond. Teachers were most likely to respond to repeated attempts (55% were responded to) and pauses after miscues (62% were responded to). Teachers were least likely to respond when the students continued reading in the text after making a miscue (only 15% were responded to). It seems reasonable to conclude from these data that there are at least three factors directly related to criteria for selection of which miscues to respond to: (1) the degree of meaning change involved; (2) the density of miscues; and (3) the strategy the student exhibits immediately following the miscue.

Timing: When did teachers respond?

The timing of teacher response was monitored in two ways: First, in terms of elapsed time between the occurrence of the miscue and the initiation of feedback; and second, in terms of the point in the text relative to the miscue at which the feedback was first offered. In general it can be said that teachers interrupt early and fast. Verbal responses were offered immediately (i.e., within 0-3 seconds), almost 75% of the time and before the student had progressed very far beyond the miscue in the text (e.g., 83% before the next sentence break).

In comparing the timing of responses between easy and hard materials it was found that point of response tended to be earlier in the more difficult material, although elasped time was greater. This phenomenon can be explained in part by the associated decrease in continuations by the students when moving from easy to difficult material (35% to 33%), and the increase in repeated attempts (16% to 18%) and pauses (7% to 8%). When the feedback is offered, then, it is directly related to the degree of text difficulty in relation to pupil ability. Where the feedback is offered is influenced by the student's strategy following the miscue.

Form: What kind of feedback was offered?

When teachers did respond overtly to student miscues, their responses were divided fairly evenly (19% versus 18%) between terminal feedback (initially giving the student the text word) and sustaining feedback (attempting to have the student identify the text word). The data also revealed that inservice teachers resorted more often to terminal feedback than preservice teachers and that both groups used significantly more terminal feedback when students were reading in the more difficult material. See Table 2.

In terms of breakdown of teacher sustaining feedback behaviors, it was found that inservice teachers relied on significant attending feedback more often than did preservice teachers (38% as opposed to 22%). Significant attending feedback provides the student with an opportunity to respond and is noncue focused. Examples would include such statements as: "try again" or "keep working at it." Both groups tended to rely less on significant attending feedback when students were reading in the more difficult material.

Both groups of teachers were fairly evenly split between their reliance on grapho-phonic and contextual prompts. As a group, teachers became more contextually oriented in their prompts as students read in the more difficult material. While, as noted earlier, teachers were more likely to respond to miscues which substantially affected text meaning, there was no apparent relationship

TABLE 2
Percentage of Teacher Prompts as a
Function of Sustaining Feedback by Category

MATERIAL TYPE	*Teachers*	*Total Number of Miscues*	*Attending Without Prompt*	*Simple Graphic/Phonic*	*Complex Graphic/Phonic*	*Simple Context*	*Complex Context*
EASY	Inservice	67	34%	31%	5%	25%	5%
	Preservice	68	28%	40%	1%	27%	1%
	Total	135	31%	36%	3%	26%	4%
HARD	Inservice	83	43%	24%	0%	22%	11%
	Preservice	77	14%	39%	0%	43%	4%
	Total	160	29%	31%	0%	32%	8%
ALL	Total	295	30%	33%	2%	29%	6%

between the level of meaning change and the form of sustaining feedback, nor was there a discernible relationship between the form of sustaining feedback and the degree of grapho-phonic similarity between the miscue and the text word. In other words, a student who generated a miscue which substantially affected text meaning was no more likely to get a context prompt than a grapho-phonic prompt. Conversely, a student who generated a miscue which had low grapho-phonic similarity to the text word was no more likely to get a grapho-phonic prompt than a context prompt. It was surprising—particularly with respect to miscues which substantially affected meaning—that teachers would apparently use this information to determine whether or not they should prompt and then not use the information in determining the kind of prompt they would offer.

Also, teachers rarely initiated a prompt at one level and completed the prompt at another. For example, in only 6% of the cases in which a prompt was offered did teachers combine both grapho-phonic and contextual cues. Teachers were more likely to make repeated prompts at the same level even when faced with unsuccessful responses by the student.

An analysis of the amount of time spent on a prompt from initiation by the teacher to resumed reading by the pupil revealed no significant differences between the groups of teachers. There was, however, a slight tendency toward shorter interruptions in the more difficult material. A significant difference ($p <$ .05) was found for the amount of time away from the task of reading relative to the form of the sustaining feedback with grapho-phonic prompts taking much longer than the others.

Of all the single word miscues made by students not given overt feedback, 45% were utlimately identified by the students themselves. Of those single word miscues responded to and receiving sustaining feedback, 41% were ultimately identified by the teacher and 54% by the students. Some forms of prompts were more associated with student identification of the text word than were others. Simple context prompts, for example, led to student identification 67% of the time, while simple grapho-phonic prompts led to student identification only 51% of the time.

SUMMARY

This study was designed in order to observe and describe the characteristics of verbal feedback behaviors used by teachers during guided oral reading. Basic relationships between pupil behavior and teacher behavior were explored.

In summary, the major findings of this study can be categorized by characterizing teachers' verbal feedback in terms of selection, timing and form. On the whole, both experienced and inexperienced teachers were remarkably similar in their choices of response patterns. With regard to selection, teachers were more likely to respond to miscues made in difficult material. As well, all teachers appeared sensitive to meaning change in determining which miscues they would respond to (although they did not appear to use this information in determining the form of their verbal feedback). With regard to the timing dimension of teacher feedback, teachers more often responded to miscues quickly rather than delaying their point of intervention. Finally, in the area of form, terminal feedback appeared to be a strategy teachers turned to more often in difficult text. Attending prompts (i.e., sustaining feedback with no cue offered), simple grapho-phonic prompts and simple contextual prompts were equally divided. The patterns of sustaining feedback seemed to indicate that contextually oriented prompts took less time and were more likely to lead to student identification of miscues than grapho-phonically oriented prompts.

The generalizability of the findings from this study are limited by the dyadic context and the unfamiliar teacher/pupil pairings in which the interactions occurred. But as Wilmot (1975) has pointed out, the basic components of a communicative system may be more easily studied initially in a dyadic setting. The results of this study form a useful and necessary basis for expanded studies of teacher-pupil interactions during oral reading in the classroom.

ACKNOWLEDGEMENT

The research reported herein was supported in part from a grant by the National Institute of Education (Contract No. G-800-033). It is part of an ongoing study into teacher verbal feedback to student miscues.

REFERENCES

Allington, R. L. *Are good and poor readers taught differently? Is that why poor readers are poor readers?* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Anderson, L. & Evertson, C. *An experimental study of effective teaching in first grade reading groups* (Tech. Rep. 4072). Austin: The Research and Development Center for Teacher Education, The University of Texas, 1978.

Artley, A. S. Oral reading as a communication process. *Reading Teacher,* 1972, *36,* 46-51.

Brady, M. E., Lynch, W. W., & Cohen, D. A. *Oral Reading Observation System—Observer's Training Manual.* Bloomington: Center for Innovation and Teaching the Handicapped, Indiana University, 1976.

Brophy, J. E. & Good, T. L. *Teacher-child dyadic interaction: A manual for coding classroom behavior.* Austin: The Research and Development Center for Teacher Education, The University of Texas, 1969.

Goodman, K. S. *Theoretically based studies of patterns of miscues in oral reading performance* (Final Report, Project No. 9-0375). Washington: U.S. Office of Education, 1973.

Howlett, N. & Weintraub, S. Instructional procedures. In R. C. Calfee & P. Drum (Eds.), *An Analytic Study of Compensatory Reading Programs.* Newark, Delaware: International Reading Association, 1978.

Hoffman, J. V. On providing feedback to reading miscues. *Reading World,* 1979, *19,* 342-350.

Hoffman, J. V. & Baker, C. *Observing communication during oral reading instruction: A critique of past research and a report on the development of a taxonomy of behaviors useful in field-based research.* Paper presented at the XXX International Conference on Communication: Acapulco, Mexico, 1980.

Hoffman, J. V., Gardner, C. & Clements, R. *FORMAS—Feedback to Oral Reading Miscus Analysis System: Training Manual.* Austin: The Research and Development Center for Teacher Education, The University of Texas, 1980.

Hoffman, J. V. & Clements R. *A descriptive study of the characteristics of miscue focused verbal interactions between teacher and student during guided oral reading.* Paper presented at the annual conference of the International Reading Association, New Orleans, May 1981.

Niles, J. *Effects of teacher feedback on beginning reader's oral reading behavior.* Paper presented at the annual meeting of the National Reading Conference, San Diego, 1980.

Stallings, J., Needels, M. & Stayrook, N. *The teaching of basic reading skills in secondary schools, Phase I and Phase III.* Menlo Park, California: SRI International, 1979.

Weber, R. First graders' use of grammatical context in reading. In H. Levin & J. P. Williams (Eds.), *Basic Studies in Reading.* New York: Basic Books, 1970.

SANDRA BERKOWITZ
BARBARA TAYLOR
University of Minnesota

THE EFFECTS OF TEXT TYPE AND FAMILIARITY ON THE NATURE OF INFORMATION RECALLED BY READERS

The purpose of this study was to examine how different structural and content variables operating within texts affect the nature of sixth-grade students' recall for reading. It was anticipated that such an investigation would improve our understanding of children's processing and memory for different text types. Classroom teachers and reading researchers have observed that students have difficulty reading and remembering expository text (Baker and Stein, 1978). One reason for this difficulty could be that the text-specific, hierarchical structure of exposition is not easy for students to discern and use as a memory framework. In light of this difficulty, students' memory for expository writing was compared with their memory for a narrative passage organized around a conventional, temporal structure.

Another factor that is frequently cited as a reason for students' difficulty with expository text is that this mode of discourse contains difficult or unfamiliar concepts. In order to examine the effect of familiarity on memory for text, an expository passage written on a familiar topic (animal protection) and an expository passage written on an unfamiliar topic (bird migration) were included. Both passages were written using precisely the same organization of main ideas and details. In order to validate the relative familiarity of the two passages, fifth-grade students at the school where the study was conducted answered questions from the two texts without reading the texts. A *t*-test revealed that the children were able to answer significantly more questions on animal protection than on bird migration ($t_{58} = 10.46$, $p < .01$). This suggests that the animal protection passage contained more familiar concepts for the children than the bird migration passage.

In terms of examining the nature of the information that the sixth-grade students recalled from these three texts, it was felt that a fine-grained procedure of recording different types of information recalled would be more informative than a global procedure of counting total number of verbatim or paraphrased idea units recalled. According to a text comprehension and production model developed by Kintsch and van Dijk (1978), readers' free recall protocols consist of four types of information: reproductions, or verbatim text information; reconstructions, or inferences integrating prior knowledge and text information; erroneous statements about text content; and metastatements about the material. Drum and Lantaff (1978) have further specified the types of information found in recall protocols to include five categories ranging from text-specific information that corresponds with text units unambiguously to text-external information which bears no relationship to the stimulus material. In order to compare students' recall protocols of different texts in this study, Drum and Lantaff's and Kintsch and van Dijk's classification systems were modified to include the following four categories: text-specific or verbatim and paraphrased units from the text, text-entailed information consisting of necessary inferences implied by the text, text-embellished units which were related to the original text but were neither stated in nor implied by the material, and intrusions consisting of erroneous or unrelated information.

The following hypotheses were made regarding students' memory for the four recall categories across the three text types under investigation. First, it was predicted that the students would be able to recall more text-specific information from the narrative passage than from either expository passage and recall more text-specific information from the familiar expository passage than from the unfamiliar expository passage. Regardless of content, the stereotypic, temporal organization of ideas in a story should facilitate memory for text-specific information to a greater extent than the less conventionalized, hierarchical organization of ideas in an expository text. With text structure controlled, the influence of prior knowledge for text content on recall of text-specific information should be greater for a familiar expository passage than for an unfamiliar expository passage.

It was also predicted that the students would be able to recall more text-entailed and more text-embellished information from the narrative and familiar expository passages than from the unfamiliar expository passage. Students' schemata for the sequence of familiar events found in the narrative passage and for the content of the familiar expository passage should enhance their abilities to make inferences and embellishments while reading and recalling these texts. On the other hand, making inferences and embellishments while reading unfamilar text should be more difficult for students due to their lack of prior knowledge of this text.

Finally, it was expected that the students' recalls would contain more intrusions for the unfamiliar expository passage than for the narrative or familiar expository passages. Students' lack of prior knowledge about the unfamiliar passage should cause them to make a relatively high proportion of errors when trying to recall the passage. This problem should not exist to the same extent with the familiar expository and narrative passages.

METHOD

Subjects: the subjects in this study consisted of sixteen competent and nineteen less competent sixth-grade readers from an urban elementary school. The competent readers scored on or above grade-level and the less competent readers scored below grade-level on the Reading Comprehension subtest of the Stanford Achievement Test.

Materials: The materials consisted of a narrative and two expository passages of approximately equal length (550-600 words), number of idea units (63) and readability (fourth grade). The expository texts were written using the same organizational structure of main ideas and details. The narrative passage was about a farmer who made repeated and unsuccessful attempts to teach a talking crow to be useful. The familiar expository passage dealt with how animals protect themselves and the unfamiliar exposition contained information about bird migration.

Procedure: Each subject read one of the three passages each week for three weeks. On the day after reading a passage, subjects were instructed to write everything they could remember about the material. Each protocol was then analyzed to determine the number of text-specific, text-entailed and text-embellished idea units as well as the number of intrusions it contained using a modification of a text analysis procedure developed by Meyer (1975). A 2 (reading ability) × 3 (text type) × 4 (recall category) ANOVA, repeated on the last two measures, was then conducted on the raw scores. An additional 2 × 3 × 4 repeated measures ANOVA was used to analyze percent scores. To obtain the percent scores, each subject's raw scores within each text type were converted to a percent of the total idea units recalled from that text. These scores reflected the proportion of total units per passage that were comprised of each recall category.

RESULTS AND CONCLUSIONS

There was no significant main effect for reading ability for raw scores, $F<1$, or for percent scores $F<1$. Furthermore, there were no significant interactions involving ability and any other variable in the analysis of raw or percent scores. The analysis of raw scores (see Table 1) revealed significant effects for text type, $F\ (2, 66) = 84.3\ p < .001$, recall category, $F\ (3, 99) = 123.3$, $p<.001$, and a significant text type x recall category interaction, $F\ (6, 198) = 61.6$, $p < .001$. Tukey's post hoc analysis of the raw scores for text types revealed that students wrote significantly more idea units for the narrative passages than for either of the expository passages $p < .05$ (1) and more information for the familiar vs. the unfamiliar expository passages (1) $p < .05$.

In light of this tendency of subjects to recall more information in general from the narrative passage vs. the familiar expository passage vs. the unfamiliar expository passage, it seemed that it would be more informative to consider percent scores than raw scores in comparisons of the three text types within the four recall categories. Analysis of the percent scores revealed a significant main effect for recall category, $F\ (3, 99) = 175.6$, $p <.001$, and a significant text type x recall interaction, $F\ (6, 198) = 11.7$, $p <.001$.

TABLE 1
Mean Raw and Percent Scores and Standard Deviations
for the Three Passages and Four Recall Categories*

	Passage Type					
	Narrative		Familiar Expository		Unfamiliar	
Recall Expository Category	Raw	Percent	Raw	Percent	Raw	Percent
Text Specific	M 20.4	71.6	8.3	59.4	6.1	56.4
	SD 8.6	14.5				
Text Entailed	.3	2.1	15.0	1.0	8.9	
	SD 2.2	10.1	2.2	13.7	1.1	10.0
Text Embellished	M 2.0	7.2	.7	7.8	.2	2.4
	SD 1.6	5.6	1.2	10.7	.6	6.1
Intrusions	M 1.7	6.8	1.0	9.7	3.0	29.5
	SD 1.9	8.7	1.6	17.8	3.0	23.0
Marginal	M 27.5		12.1		10.3	

*n = 35

Tukey's post hoc analyses revealed that proportionately more text-specific idea units were recalled for the narrative passage than either expository passage ($p<.05$). This finding is consistent with the hypothesis that the stereotypic temporal structure of narrative material is easier for students to detect and use to facilitate memory for text-specific information than the less conventionalized hierarchical structure of exposition. There was no difference in the proportions of text-specific idea units recalled from the two expository texts. Familiarity of content did not appear to affect the proportions of recall from this category of information.

For the proportions of text-entailed and text-embellished information, no significant differences were found between the three passages. This finding is of particular interest in a comparison of the two expository texts. Given most theories of discourse processing, especially those emphasizing the power of schemata to facilitate inferencing, one would expect to find more text-entailed and text-embellished information produced from stimulus material written on a familiar topic than from text written on an unfamiliar topic (Adams and Collins, 1977; Rumelhart and Ortony, 1977). On the other hand, two explanations could account for a finding of no difference. First, readers may edit out available information from prior knowledge at the time of recall because they are aware that this information was not present in the text. Second, as suggested by Spiro and Taylor (in press), students may compartmentalize information learned in school environments instead of integrating it with existing knowledge.

Subjects in this study produced proportionately more erroneos or unrelated idea units for the unfamiliar exposition than for either of the other texts. This finding appears to support the hypothesis that lack of prior knowledge of a topic may cause readers to confuse information while reading or to access unrelated information while attempting to recall material that was never clearly understood.

Implications for the classroom based on these findings are as follows. First, students appear to be comparatively good at recalling text-specific information from narrative material. Perhaps increased instructional time should be devoted to helping students read and remember expository text as they appear to be less proficient with this mode of discourse. Additionally, specific instructions in discerning the structure of expository writing and using that structure as a framework for memory may increase students' ability to recall informative reading. When assigning expository reading on an unfamiliar topic, time spent introducing unfamiliar concepts and building background before reading and reviewing the material after reading may help students decrease the number of intrusions they produce. Finally, the finding that there was not a significant main effect for reading ability suggests that poorer readers can perform similarly to more competent readers if the material is written at below-grade readability levels.

REFERENCES

Adams, M. and Collins, A. *A schema-theoretic view of reading comprehension.* (Tech. Rep. No. 32). Champaign, Illinois: Center for the Study of Reading, University of Illinois, April 1977. (ERIC Document Reproduction Service No. ED 144039)

Baker, L. and Stein, N. *The development of prose comprehension skills.* (Tech. Rep. No. 102). Champaign, Illinois: Center for the Study of Reading, University of Illinois, April 1977. (ERIC Document Reproduction Service No. ED 159 663)

Drum, P. and Lantaff, R. *Scoring cagegories for protocols*. Unpublished manuscript, University of California, Santa Barbara, 1978.

Kintsch, W. and van Dijk, T. Toward a model of text comprehension and production *Psychological Review*, 1978, *85*, 363-394.

Meyer, B. *Organization of prose and its effect on recall*. Amsterdam: North Holland Publishing Company, 1975.

Rumelhart, D. and Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, and W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Erlbaum, 1977.

Spiro, R. and Taylor, B. On investigating children's transition from narrative to expository discourse: toward a multi-dimensional psychological classification of texts. In R. Tierney, J. Mitchell, and P. Anders (Eds.), *Understanding readers' understanding*. Hillsdale, N.J.: Erlbaum, forthcoming.

LEA M. MCGEE
Louisiana State University

GOOD AND POOR READERS' ABILITY TO DISTINGUISH AMONG AND RECALL IDEAS ON DIFFERENT LEVELS OF IMPORTANCE

Adults and children who are fluent readers recall the more important information in a passage better than they recall the less important information (Johnson, 1970; Meyer & McConkie, 1973; Thorndyke, 1977; Brown & Smiley, 1977; Waters, 1978; Taylor, 1980). In contrast, less mature readers do not give prominence to recall of the more important information (Smiley, Oakley, Worthen, Campione, & Brown, 1977; Eamon, 1978-79; Taylor, 1980). Yet, the ability to note and recall important ideas in a passage is an essential reading ability which all readers should have facility using. In order to provide insight into techniques which might improve this ability in less mature readers, the purpose of the present study was to examine processes and factors which may influence recall of important information.

Differential recall of important information may be due to selective processing of that information (Eamon, 1978-79). Fluent readers do not process every letter or word as they read; they process only that information which is most relevant to comprehending the text (Goodman, 1970; Gibson & Levin, 1975; Walker, 1975-76). Therefore, fluent readers may recall more important ideas better than less important ideas because they select and more thoroughly process these ideas. Conversely, less mature readers may recall less of the important ideas because they do not select and more thoroughly process these ideas.

An assumption in the selective processing of important information notion is that fluent readers evaluate the importance of ideas so that they can select important ideas for more thorough processing (Ashley & Fritzen, 1980). One reason that less mature readers do not have superior recall of more important information may be that they do not adequately evaluate the importance of ideas.

There is only partial support for the evaluation difference notion in previous research. Eamon (1978-79) found good adult readers better able to evaluate text importance than poor readers, but Ashley and Fritzen (1980) did not find differences for adult good and poor readers in this ability. The one study which examined the ability of young readers to determine the importance of text information demonstrated that third graders do not distinguish among ideas on four levels of importance while fifth graders only separate the top level ideas from ideas on other levels of importance (Brown & Smiley, 1977). There is a lack of research examining this ability in young good and poor readers. Therefore, one purpose of the present study was to provide more conclusive evidence about good and poor readers' ability to evaluate the importance of text information.

One methodological consideration given attention in the present study was the level of difficulty of the text in relation to the comprehension ability of the readers. In previous studies (Eamon, 1978-79; Brown & Smiley, 1977) comprehension difficulties may have prevented readers from adequately assessing the importance of the content. Adult poor readers in Eamon's study may not have had sufficient reading ability to comprehend the collegel level passages. Similarly, third graders in Brown and Smiley's study may not have had adequate listening abilities to fully comprehend the fifth grade passage. The present study employed passages written with readability levels similar to the comprehension level of the readers. Moreover, good and poor readers with similar comprehension abilities were selected. An additional group of readers for whom the passage was very easy to comprehend was included in order to investigate how text difficulty affects ability to evaluate importance.

EXPERIMENT 1: METHODS

One hundred twenty third and fifth grade students from four elementary schools participated in the study. Forty third grade good readers were randomly selected from all third graders comprehending on a 3.0-4.5 grade level (SRA, Primary II). Forty fifth grade poor readers were randomly selected from all fifth graders comprehending on a 3.0-4.5 grade level (SRA, Multilevel Blue). Forty fifth grade good readers were randomly selected from all fifth graders comprehending on a 5.5-7.0 grade level (SRA, Multilevel Blue). Twenty in each of the three reading ability/grade groups were randomly selected to participate in Experiment 1 and the remaining twenty in each group participated in Experiment 2.

Two 125-word expository passages written by the researcher were segmented into idea units, and these units were placed in a hierarchy of importance using an analysis of discourse procedure (Taylor, 1978) adapted from Grimes (1975) and Meyer (1975). The two passages (presented in figure 1) were written with a third grade readability level (Fry, 1977), an identical three-level hierarchy of importance, and a similar number of idea units (61 and 62).

Figure 1
Experimental Passages

SEEDS

All kinds of seeds are carried to new homes in different ways. Some seeds are carried to their new homes by the wind. The wind can blow dandelion seeds across the grass to some other yard. If the wind blows other seeds to a good spot with lots of sun and water, new trees or a new plant will grow. Seeds can be carried to their new homes by furry animals or people. Some seeds may hang in animals' fur. Or seeds may stick in their paws. Some seeds may catch in peoples' clothes. Birds can carry seeds and nuts to new homes. If birds drop seeds or nuts they like to eat on rocks, then a new plant or new tree will not grow there.

ANIMAL TEETH

Many kinds of animals use their teeth in different ways. Some animals use their front teeth to cut things. Lots of beavers cut down trees with their strong front teeth. Rabbits use these teeth to cut their food like carrots. Animals use other teeth to tear things. When an animal like a bear or a wolf wants to tear its food, it uses the teeth at the side of its mouth with the points. Dogs and cats also use these teeth to tear things. Animals use still other teeth to chew their food. When cows chew the grass they eat, they use the teeth in the back of their mouths. Horses also use these teeth to chew the hay that they eat.

Two rating statements were created for each of the three levels of importance by randomly selecting sentences from each level in the hierarchy. Level 1 statements were selected from six sentences at the lowest level of importance.

Level 2 statements were selected from three sentences at the middle level of importance. Level 3 statements were the one sentence at the highest level of importance and a matching sentence rewritten in the passive voice.

A booklet containing a practice passage, two practice rating statements, the experimental passages, and experimental rating statements was used in the testing procedure. Order of passages was counterbalanced across booklets, and order of rating statements was randomized.

Each student read the practice passage. Then the researcher demonstrated the rating procedure. The students' attention was directed to the five circles of graduated sizes which followed each rating statement. The student was told that some ideas in a passage were more important than other ideas. The circles represented how important sentences were in the passage. The most important sentences were represented by the largest circle. These sentences were the big ideas of the passage or the main idea. The least important sentences in the passage were represented by the smallest circle. These sentences were the little ideas of the passage or the details. The researcher indicated that the first practice rating statement was the most important idea in the passage, and the largest circle should be marked. The student was told that the second statement was not an important idea, it was the least important idea. The smallest circle should be marked. The student was reminded that the bigger the circle, the more important the idea. An additional comment was made that some ideas were not the most important or the least important but were in between. The student was told to mark the circle that showed how important the ideas were according to the size of the circle. Then, each student was directed to read the passage, complete a distractor addition problem (to eliminate short term memory), and rate the six statements. Students completed the passages and rating statements at their own pace.

Rating statements which received a mark in the largest circle were assigned 5 points in the scoring procedure. A mark in the next largest circle was assigned 4 points, and so on.

RESULTS AND DISCUSSION

After collapsing across passages a 3 (reading ability/grade) × 3 (level of importance) mixed analysis of variance was conducted. Reading ability/grade was a between-subjects variable, and level of importance was a within-subjects variable.

Level of importance was highly significant, $F\ (2,\ 51) = 157.80.\ p < .001$. Duncan's multiple range tests at the .05 significance level revealed that level 3 statements ($\bar{X} = 4.33$) were judged to be more important than level 2 ($\bar{X} = 3.0$), and these statements were judged to be more important than level 1 statements ($\bar{X} = 2.43$).

The reading ability/grade factor was not significant. Therefore, all three reading ability/grade groups rated the highest level ideas as being more important than the middle level ideas, and those ideas were rated as being more important than the lowest level ideas. There was a significant interaction between the reading ability/grade and level of importance factors, $F\ (4, 51) = 8.15, p < .001$. As shown in figure 2, fifth grade good readers rated level 3 ideas as being significantly more important than either fifth grade poor readers or third grade good readers.

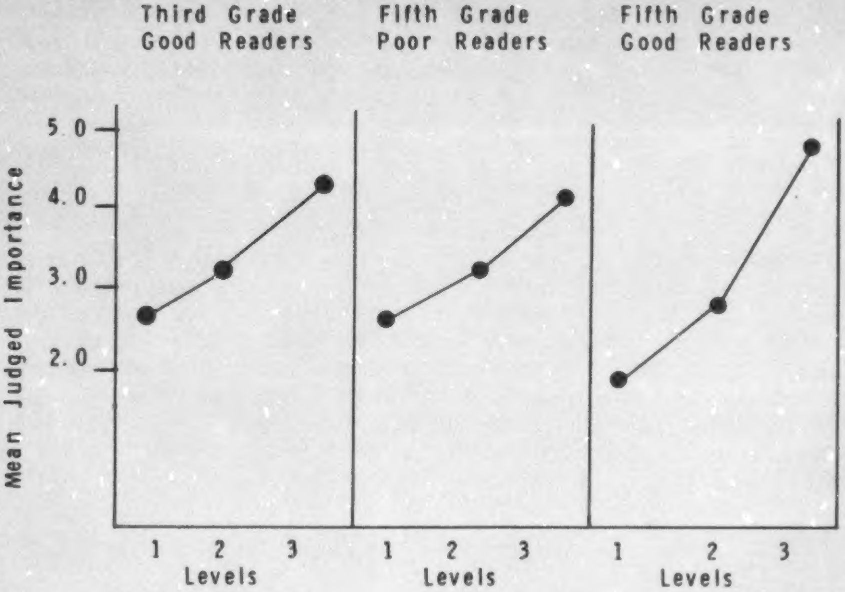


Figure 2. Judged importance of statements on three levels of importance.

Similarly, this figure illustrates that fifth grade good readers judged level 1 ideas to be significantly less important than fifth grade poor readers or third grade good readers. Fifth grade good readers made a greater distinction between the most important and least important information in the passages than the third grade good or fifth grade poor readers.

These results indicate that poor readers (grade 5), when compared to good readers (grade 3) who are reading on similar grade levels, are equally as sensitive to importance of ideas as good readers. Good and poor readers of different grade placements may judge the importance of ideas equally well because they are evaluating information of similar difficulty. However, poor readers, when compared to good readers of the same grade placement but not on similar reading levels, are not as sensitive to importance of ideas as good readers. Good readers may evaluate the importance of text information better than poor readers of the same grade placement because they are judging relatively easier reading material. Hence, good readers may not evaluate the importance of text information any better than poor readers unless good readers are given significantly easier text from which to evaluate importance than poor readers.

EXPERIMENT 2

Students' ability to rate the importance of text has been suggested to influence recall of information on different levels of importance (Eamon, 1978-79). Therefore, good and poor readers' recall of ideas on different levels of importance should reflect how well they evaluate the importance of passage information. Since fifth grade good readers make a greater distinction between the most important and the least important ideas in a passage than fifth grade poor readers or third grade good readers, they were expected to recall the most important ideas better than members of the other reading ability/grade groups.

METHODS

A booklet containing a practice passage and two experimental passages identical to those used in experiment 1 was presented to each of the sixty students as they arrived for testing. Each student was directed to read the practice passage, complete a distractor addition problem, and then was asked to tell everything he/she could remember about the passage. This procedure was repeated for each experimental passage. Recalls were tape recorded.

Recalls were scored for proportion of idea units retold at each level of importance by two independent raters. Interrater reliability between these raters was .98.

RESULTS AND DISCUSSION

After collapsing across passages a 3 (reading ability/grade) x 3 (level of importance) mixed analysis of variance was conducted. Reading ability/grade was a between-subjects variable, and level of importance was a within-subjects varaible.

The reading ability/grade factor was significant, $F\ (2, 46) = 100.46$, $p < .001$. Duncan's tests revealed that fifth grade good readers recalled a significantly greater proportion of idea units ($\bar{X} = .42$) than fifth grade poor readers ($\bar{X} = .27$) who recalled a greater proportion of idea units than third grade good readers ($\bar{X} = .15$).

The level of importance factor was not significant but the expected interaction between reading ability/grade and level of importance factors was significant, $F\ (8, 46) = 9.12$, $p < .001$. Figure 3 illustrates that fifth grade good readers recalled proportionately more level 3 ideas than either fifth grade poor or third grade good readers. This figure also shows that only fifth grade good readers recalled more level 3 ideas than level 2 or 1 ideas.

These results suggest that the ability to evaluate the importance of ideas is reflected in recall of those ideas. Fifth grade good readers not only are better able to distinguish the most important ideas from other ideas but are also better able to recall those ideas.

GENERAL DISCUSSION

The purpose of the present study was to examine good and poor readers' ability to evaluate the importance of text and to investigate the influence of this ability on recall of important information. The results indicate that fifth grade good readers are better able to distinguish the importance of information than fifth grade poor or third grade good readers. However, these results do not support the

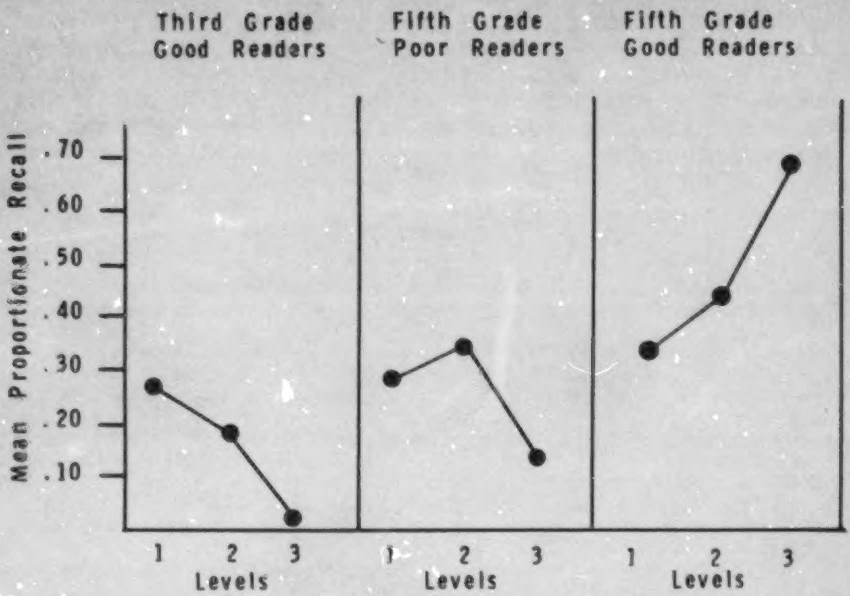


Figure 3. Proportionate recall of text information on three levels of importance.

conclusion that poor readers are insensitive to importance. Fifth grade poor readers, as well as third grade good readers, were able to judge the important and unimportant information. All readers in the present study were able to make better distinctions among ideas on diffent levels of importance than students of similar ages were able to make in a previous study (Brown & Smiley, 1977). Students in the present study may have been able to make these superior judgments due to differences in the materials used in the two studies. Students in the present study may have benefited from reading relatively easier text, shorter passages, and material about more familiar topics. Certainly the performance of the fifth grade good readers indicates that reading easy text influences how well students judge the importance of text.

The results of the present study may support the notion that important ideas are better remembered because they are selected for more thorough processing. Fifth grade good readers may have more thoroughly processed the important ideas; they were better able to distinguish the more important ideas and recalled more of these ideas than fifth grade poor and third grade good readers. However, these results do not suggest that poor readers are not selective processors because they do not adequately evaluate the importance of information. Poor readers were able to accurately judge important and unimportant information. Moreover, the results may not suggest that good readers are more selective processors than poor readers. Poor readers, when compared to good readers of similar reading level, did not recall or judge the importance of ideas differently than good readers. However, difficulty of the text seems to influence ability to recall and judge the importance of ideas. The present study only examined how good readers performed when given relatively easy material to read. Further research giving both good and poor readers relatively easy material is needed before a conclusion can be drawn whether there are differences among good and poor readers' selective processing of ideas. Finally, further research is needed to explore why third grade good and fifth grade poor readers did not recall the more important ideas better than the less important ideas when they were able to judge the importance of these ideas.

REFERENCES

Ashley, S. & Fritzen, J. *Good and poor reader evaluation and comprehension of prose passages.* Paper Presented at the American Educational Research Association Conference, Boston, MA 1980.

Brown, A. & Smiley, S. Rating the importance of structural units of prose passages: a problem of metacognitive development. *Child Development*, 1977, *48*, 1-8.

Eamon, D. Selection and recall of topical information in prose by better and poorer readers. *Reading Research Quarterly*, 1978-79, *14*, 244-257.

Fry, E. Fry's readability graph: clarifications, validity, and extention to level 17. *Journal of Reading*, 1977, *21*, 242-252.

Gibson, E. & Levin, H. *The Psychology of Reading*, Cambridge, MA: M.I.T. Press, 1975.

Goodman, K. Reading: a psycholinguistic guessing game. In H. Singer & R. Ruddell (Eds.) *Theoretical Models and Processes of Reading* (2nd Ed). Newark, DE: International Reading Association, 1980, 497-508.

Grimes, J. *The Thread of Discourse*. The Hague, Netherlands: Mouton and Company, 1975.

Johnson, R. Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 12-20.

Meyer, B. *Organization of Prose and its Effects on Recall.* Amsterdam: North Holland Publishing Company, 1975.

Meyer, B. & McConkie, G. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, *65*, 109-117.

Smiley, S., Oakley, D. Worthen, D. Campione, J., & Brown, A. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Journal of Educational Psychology*, 1977, *69*, 381-387.

Taylor, B. *A Comparison of Good and Poor Readers' Ability to Read for Meaning*. Unpublished dissertation, Virginia Polytechnic Institute and State University, 1978.

Taylor, B. Children's memory for expository text after reading. *Reading Research Quarterly*, 1980, *15*, 399-411.

Thorndyke, P. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

Walker, L. Comprehending written and spontaneous speech. *Reading Research Quarterly*, 1975-76, *11*, 144-167.

Waters, H. Superordinate-subordinate structure in semantic memory: the roles of comprehension and retrieval processes. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 587-597.

BARBARA J. RENNIE
ALLAN R. NEILSEN
CARL BRAUN
University of Calgary

THE EFFECTS OF TYPOGRAPHICAL CUEING ON MEMORY FOR SUPERORDINATE STRUCTURES IN CONNECTED DISCOURSE

This study had two purposes; first, to investigate the effect of typographical cueing on good and poor readers' comprehension of and memory for familiar and unfamiliar text; second, to assess good and poor readers' levels of metacognitive awareness of their processing of text.

The theoretical framework and rationale for this study drew from three domains: schema theory, metacognition, and typographical cueing. Schema-theoretic research has demonstrated both the psychological reality of macro-structures in the reading of exposition (Meyer, 1975; Meyer, Brandt, & Bluth, 1978), as well as the influence of prior knowledge on comprehension and recall of information in text (Anderson, Reynolds, Schallert & Goetz, 1976).

In metacognitive research, attempts have been made to explore readers' introspective judgements about text processing, particularly their metacomprehensive and metamemorial awareness. Successful identification of top-level structures has been found to be a developmental skill, with accuracy generally improving with maturity (Brown & Smiley, 1977a, 1977b). Other work has revealed inconsistencies between the subjects' awareness of the relative efficiency of processing techniques, their actual processing procedures, and their explicitly-stated views of their processing (Brown, 1977).

Research related to the effect of typographical cueing has shown that cues (eg. bold face type, underlining) provide an isolation effect by visually setting apart those words cued, and that this distinction results in a higher probability of recall of cued information (Wallace, 1965). Cueing that highlights high-level, integrative text propositions has been found most beneficial to recall (Ausubel, Novak & Hanesian, 1978; Cashen & Leicht, 1970; Rickards & August, 1975).

DESIGN

Three separate experiments were carried out to meet the purposes of this study. In all three, the independent variables were reading achievement, content familiarity, and typographical cueing. Good (stanines 7-9) and Poor (stanines 1-4) achievement groups were established on the basis of Gates-MacGinitie (1978) scores. Familiar (Driver's License, Energy Sources) and Unfamiliar (Continental Drift and Navigation) passages were obtained from a previous study (McKenzie, 1980) in which a series of procedures (word association, wh-probes, student-ranking, cloze) were used to determine differences in content familiarity. The effect of typographical cueing was assessed by the systematic inclusion or exclusion of italics to cue the superordinate structure (i.e. main ideas and supporting details) of each paragraph of the passages. In addition, readability, content structures and passage order were controlled.

Subjects

A sample of 24 good and 24 poor readers was selected from seventh grade students to participate in Experiments I and II. A representative sub-sample of 12 was selected to be interviewed in Experiment III.

Procedures

In Experiments I and II, good and poor achievers were randomly assigned to each cueing condition. Half of each achievement group received Cued material while the other half received Uncued material. Good Achievers received grade 7 readability level passages while Poor Achievers received grade 4 readability level passages. Subjects were given prior instruction regarding the meaning of *topic, main idea, paragraph,* and the procedure to be followed in the task completion.

In Experiment II, subjects were required to do two application tasks. The first involved locating and stating the main idea of each paragraph in the passage. The second involved writing three questions based on "the most important" information in the passage as a whole.

Responses for Experiments I and II were scored by the main investigator and two independent scorers (interscorer reliabilities were .98 and .99).

In Experiment III, individual tape-recorded interviews were conducted to determine subjects' awareness of superordinate structures in text, typographical cueing, and the effects of content familiarity on comprehension and recall. Passages used in Experiments I and II were used as reference points in the interviews. All interviews were conducted on the day immediately following the completion of Experiment II and were transcribed and responses scored according to a list of target questions.

The results of Experiment I were analyzed by a three-way analysis of variance with repeated measures. Experiments II and III were analyzed descriptively.

RESULTS AND DISCUSSION

The results and interpretations are presented for each experiment in this study. Comparisons across experiments are made where appropriate.

Experiment I

The purpose of this experiment was to assess the effects of the independent variables on memory for superordinate structures. Findings indicated that high reading achievement and familiar content facilitated performance but that cueing had no effect. This was qualified by a significant ($F = 5.49$, $p < .05$) three-way interaction which showed that Poor Achievers benefited from Cueing in Unfamiliar text but not in Familiar text. It seems that cueing might make important information in relatively unfamiliar material more salient for poor readers. This increased salience might help to compensate for lack of knowledge of superordinate structures in text—something good readers seem to have available. This suggests that the effects of achievement, content familiarity and cueing are not unique and should be considered in relation to one another rather than in isolation.

Experiment II

This experiment required the subjects first to find and write the main idea from each paragraph, and second to make up and write three questions that asked about the most important information in each selection. Responses given to the second task should have been main ideas from the paragraphs which the subjects had just analyzed in the first task. Subjects' success in these tasks might be expected to reflect their performance in Experiment I since there they had been told to expect questions on the most important information. However, the majority (i.e. 70%) of subject-generated questions were at level three or lower in the content structure (Meyer, 1975). Descriptive analysis showed that while achievement was again a factor, content familiarity and cueing did not assist in the application tasks. Possible explanations for these findings are suggested in the discussion of Experiment III.

Experiment III

This experiment was designed to provide information about the metacognitive awareness of selected subjects in three areas which are detailed below. In contrast [illegible] the first two experiments, responses had little relationship to the subjects' achievement group. However, both content familiarity and cueing were identified by the subjects as facilitators of comprehension. With the exception of the main idea section, there was little relationship between expressed knowledge about any one topic and performance on related tasks in Experiments I and II. The results support previous findings in metacognitive studies that the subjects are not always able to verbalize what they do, do not always do what they say they do, and do not always do what they say should be done.

Main idea. When asked directly, less than half of the 12 subjects expressed an awareness of main idea statements and of their possible location in a paragraph. This is consistent with the findings in Experiment II where only 44% of main idea statements written were correct.

Typographical cueing. Subjects were shown uncued and cued versions of the same passage and asked which would be easier to outline. All 12 chose the cued version, with eleven of them identifying the cueing as their reason. This contradicts the findings of Experiments I and II related to cueing. However, it is important to note that the subjects were asked only to estimate which *would* be easier to outline. Actually requiring subjects to do the outlines would be the acid test of the effect of cueing. Nevertheless, it is important to know that subjects seem to be aware that typographical cues *might* be helpful in such a task.

Familiarity. There were three tasks in this section. First, subjects were given two article topics, one very familiar and one quite unfamiliar, and asked which would be easier to understand. All 12 chose the more familiar because they "already knew something about it," suggesting these subjects are aware of the advantage of familiarity. This reinforces the findings from Experiment I.

The second question asked what made a passage difficult to understand. Responses were varied, with five naming difficult vocabulary and three a lack of knowledge of the topic.

The third asked how the subjects knew if they understood. Seven responses were related to being "able to remember and think about it." Three subjects could not respond, although they said they would be able to tell if they had understood. Difficulty verbalizing in this and the previous question suggests a lack of metacognitive awareness of these factors.

IMPLICATIONS FOR FURTHER RESEARCH

The effect found for cueing in unfamiliar material could be a source for further investigation. This study indicated that Poor seventh grade achievers were better able to recall the superordinate structures in Cued Unfamiliar content (compared with recall from Uncued Unfamiliar text). It would be interesting to see if this finding is consistent in grades above and below the seventh grade, and if it ever extends to Good achievers. These findings would provide knowledge regarding the optimum time for instruction in the use of typographical cues.

Another area for further investigation is the effect of cueing at various levels in the content structure. If lower level ideas were cued, this might negate the effect of superordinacy on recall and any difference between recall from cued and uncued text would be more directly attributable to the cueing. This type of research would be even more informative if the subjects' recall of the particular level cued had been previously established.

A third area suggested for study concerns the questions asked by teachers. In this study, the predominance of students' questions at level three or lower in the content structure may be a reflection of the type of question they are accustomed to answering in the classroom. This raises concerns regarding the levels in the content structure at which teachers' questions are directed. Further, is there congruity between the questions teachers say they ask and those questions they actually give to their students; if the teacher says he asks for main idea and the first level of supporting detail, do his questions confirm that? Essentially, this would involve an investigation of the metacognitive awareness of teachers: Do teachers do as they say they do?

Finally, a larger scale study across several grades might be conducted to examine more extensively developmental changes in metacognitive awareness. This would be especially useful if it traced the use of the study skills usually presented in the upper elementary and junior high school years.

REFERENCES

Anderson, R. C., Reynolds, R. E., Schallert, D. & Goetz, E. *Frameworks for comprehending discourse* (Tech. Rep. 12). Urbana, Illinois: University of Illinois Laboratory for Cognitive Studies in Education, July, 1976.

Ausubel, D. P., Novak, J. D., & Hanesian, H. *Educational Psychology: A Cognitive View,* (2nd ed.). Toronto: Holt, Rinehart and Winston, 1978.

Brown, A. L. *Knowing when, where, and how to remember: A problem of metacognition* (Tech. Rep. No. 47). Urbana, Illinois: University of Illinois Laboratory for Cognitive Studies in Education, 1977.

Brown, A. L., & Smiley, S. *The development of strategies for studying prose passages* (Tech. Rep. No. 66). Urbana, Illinois: University of Illinois Laboratory for Cognitive Studies in Education, 1977. (a)

Brown, A. L. & Smiley, S. Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development,* 1977, *48,* 1-8. (b)

Cashen, V. M., & Leicht, K. Role of the isolation effect in a normal educational setting. *Journal of Educational Psychology,* 1970, *61,* 484-486.

Fry, E. Fry's readability graph: Clarification, validity, and extension to level 17. *Journal of Reading,* 1977, *21,* 242-252.

Meyer, B. J. F. *The organization of prose and its effects on memory.* New York: North Holland Publishing Company Incorporated, 1975.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. *Use of author's textual schema: Key for ninth graders' comprehension.* Paper presented at the American Educational Research Association Convention in Toronto, Canada, March, 1978.

McKenzie, G. G. *The effects of linguistic connectives and prior knowledge on comprehension of good and poor readers*. Unpublished master's thesis, University of Calgary, 1980.

Neilsen, A. R. *The role of macrostructures and linguistic connectives in comprehending familiar and unfamiliar written discourse*. Unpublished doctoral dissertation. University of Minnesota, 1977.

Pearson, P. D., Hansen, J., & Gordon, C. *The effect of background knowledge on young children's comprehension of explicit and implicit information*. (Tech. Rep. 116). Urbana, Illinois: Center for the Study of Reading, March, 1979.

Richards, J. P., & August, G. J. Generative underlining strategies in prose recall. *Journal of Educational Psychology*, 1975, *67*, 860-865.

Wallace, W. P. Review of the historical, empirical, and theoretical status of the Von Restorff phenomenon. *Psychological Bulletin*, 1965, *63*, 410-424.

CANDACE S. BOS
University of Arizona

REMEMBERING INFORMATION FROM TEXT: A COMPARISON OF THE TEXT STRUCTURE OF EDUCABLE MENTALLY RETARDED AND AVERAGE STUDENTS' RECALLS

The fact that educable mentally retarded (EMR) students have difficulty learning to read beyond the third grade level and, for the most part, do not achieve in reading beyond the intermediate (third to fifth grade) level is well documented (Cegelka and Cegelka, 1970; Kirk, Kleibhan, and Lerner, 1978). Despite the importance of reading for school-related tasks as well as successful job performance, the characteristics of EMR students reading at the intermediate level have not been delineated. This has been due, in part, to the lack of adequate methods for measuring the reading process, particularly the comprehension of expository text. In the last five years, however, a technology has been developed which allows researchers to study reading comprehension of discourse using systematic, objective methods (Frederiksen, 1975, 1979; Kintsch, 1974; Meyer, 1975). This study utilizes discourse analysis to compare the comprehension of expository text by EMR and average students reading at the intermediate level. Specifically it focuses on the reader's sensitivity to text structure when recalling expository discourse.

Previous research has demonstrated that EMR students read commensurate with their mental-age grade-expectancy at the primary reading level (Bennett, 1932; Blake, Aaron, and Westbrook, 1969). At this level major focus is placed on decoding skills. At the intermediate reading level, however, EMR students achieve below their mental-age grade-expectancy and below the reading levels of average students of similar mental ages (Bliesmer, 1954; Sheperd, 1967; Blake et al., 1969). This difficulty in overall reading achievement at the intermediate level appears to be directly related to reading comprehension (Bliesmer, 1954; Shotick, 1960; Blake et al., 1969). For example, Blake et al. (1969) found that EMR students reading at the intermediate level scored significantly lower than average students on two-thirds of the 15 comprehension skills assessed. Whenever the material was presented in expository structure, EMR students scored lower.

Research focusing on good and poor adult readers as well as readers at various grade levels indicates that as the reader matures, sensitivity to the structure of expository text becomes more pronounced, thus facilitating comprehension (Meyer, 1975, 1977; Eamon, 1978-79; Meyer, Brandt and Bluth, 1980; Taylor, 1980). Results have demonstrated that comprehenders who utilize the author's top-level structure recall more information from the passage. In other words when a recall reflects the same organization as the passage more information is recalled.

Although reading comprehension has been pinpointed as an area of concern for EMR students reading at the intermediate level, research focusing on sensitivity to text structure is nonexistent. Therefore, this study addresses this aspect of the reading comprehension process. Specifically it predicts that EMR students should be less adept than average students of the same reading level at utilizing the author's top-level structure in recalling expository text and should recall significantly less explicit information than average readers.

METHOD

Subjects

This study compared 16 junior-high level EMR students and 16 average third-grade students of the same reading comprehension level using a two group, independent sample design. The EMR students were identified as mentally retarded, were receiving special education services and had IQ's ranging from 50 to 70 as measured on individual intelligence tests. The average students were enrolled in regular third grade classrooms, were not receiving special education or remedial reading services and had intellectual ability in the average range as judged by teachers. There was no significant difference between the two groups (t = .13) on the reading comprehension level as measured by the Stanford Achievement Test, Primary Level II, Reading Comprehension (Madden et al., 1972). The two groups were also equated on recognition of the words in the passage and background knowledge for the content of the passage. The background knowledge was measured by asking each subject to orally respond to six questions concerning beavers. Although the two groups had different chronological ages, they had similar estimated mental ages.

Materials

A 254 word expository passage on the topic of beavers was adapted from third grade reading material. The passage consisted of information concerning beavers and focused on three superordinate concepts: (a) what beavers look like, (b) why beavers live in ponds, and (c) how beavers build dams. The first concept was implicit in the text while the last two concepts were explicit.

The passage was first analyzed using an adaptation of Frederiksen's (1975) discourse analysis (Bos, 1979). Both propositions and interpropositional relations were determined with an adaptation of Meyer's (1975) discourse analysis specifying the superordinate concepts' propositions) and the placement of the subordinate propositions and interpropositional relations beneath the superordinate concepts. Two levels of concepts were specified, superordinate and subordinate. As a result three hierarchical structures were developed underneath the three superordinate concepts, the Looks Hierarchy, the Live in Ponds Hierarchy and the Build Dams Hierarchy.

In addition, twelve questions were designed to determine if the students could generate inferences left implicit in the passage. Three questions addressed the three superordinate concepts, i.e. What do beavers look like? Why do beavers live in ponds? How do beavers build dams?

Procedure

Each student worked individually with the investigator and the session was taped recorded. The subject was given a practice passage to become familiar with the procedure. The subject then read the expository passage silently and worked on an unrelated task consisting of visual closure activities. Next the subject orally recalled the passage. Finally the subject orally answered the inference questions.

Analyzing and Scoring the Recalls and Inference Questions

First, each subject's recall was analyzed using the same adaptation of Frederiksen's (1975) discourse analysis as was used for the passage. Each recalled proposition and interpropositional relation was scored as explicit or implicit by comparing the recall message base to the passage message base. Next each explicit proposition and interpropositional relation was labeled in relation to its position in the hierarchical structure of the passage message base. Finally, each subject's response to the three questions concerning the superordinate concepts was scored.

Interscorer reliability for four randomly selected recalls was .92 and for the three inference questions was .96.

RESULTS

The recalls as well as the inference questions were analyzed to determine if: (a) EMR students recall less explicit information (propositions plus interpropositional relations) than average students, (b) EMR students differ from average students in their recall of each of the three hierarchical structures, (c) fewer EMR students recall the two explicit superordinate concepts, and (d) EMR students generate fewer acceptable responses to the inference questions.

A t-test revealed that EMR students recalled significantly less explicit information contained in the three hierarchical structures than average students [t (30) = 3.43, $p<.005$]. While the mean amount of information recalled by EMR students was 6.75, average students recalled almost twice as much information, 12.5.

A 2(group) × 3(hierarchical structure) analysis of variance with a repeated measure on the second factor was used to determine if EMR students differ from average students in their recall of the three hierarchical text sturctures. The amount of subordinate information recalled was used as the measure for the recall of the hierarchical structures. The results of the main effect for group indicate that EMR students recalled less subordinate information than average students [F (1,30) = 14.86, $p<.001$]. However, there was also a significant effect for type of hierarchical structure [F (2,60) = 8.06, $p<.001$] and a significant interaction effect [F (2,60) = 11.88, $p<.001$]. The results of post hoc comparisons indicate that there were no significant differences between the two groups for the Looks Hierarchy [F (1,90) = 1.72] or the Live in Ponds Hierarchy [F (1,90) = .16]. EMR students did, however, recall significantly less subordinate information than average students concerning "how beavers build dams" [F (1,90) = 37.90, $p < .001$].

Chi-square analyses were used to determine if fewer EMR students recalled the two explicitly stated superordinate concepts. These results revealed that while there was no difference between the two groups for the Live in Ponds Concept [X^2 (1) = 1.41], there was a significant difference between the two groups at the .1 level for the Build Dams Concept [X^2 (1) = 3.79, $p<.1$].

The results of the three inference questions are in agreement with the results from the recall. While there were no differences between the two groups for the Looks Concept [t(30) = .97] or the Live in Ponds Concept [t(30) = .38], EMR students scored significantly lower than average students on the Build Dams Concept [t(30) = 2.49, $p<.01$].

DISCUSSION

The finding that EMR students have greater difficulty recalling information presented in an expository format than average students of the same reading level is congruent with other research conducted with EMR students reading at the intermediate level (Bliesmer, 1954; Shotick, 1960; Blake et al., 1969; Bos, 1979). In addition to this general finding this study focused on a specific aspect of reading comprehension, namely the EMR students' sensitivity to the text structure. Educable mentally retarded students performed equivalently with average students on hierarchical structures that required mainly conjunctive, causal and contrastive relations, i.e. the Looks Hierarchy and the Live in Ponds Hierarchy. These results imply that EMR and average students have similar comprehension of text hierarchies that utilize these types of interpropositional relations even though some of the relations had been left implicit in this particular text.

In contrast, EMR students appeared to have particular difficulty with the text hierarchy that required mainly temporal relations, the Build Dams Hierarchy. This difficulty in dealing with the temporal text hierarchy may be unique to this particular text suggesting a need for further research concerning EMR students' sensitivity to different text hierarchies. Further research should control for such text aspects as embededness and number of propositions and interpropositional relations in each hierarchy. On the other hand, EMR students may, indeed, demonstrate particular difficulty relating and remembering sequential information. Research in the area of problem solving with performance task support this notion (Spitz and Borys, 1977; Brynes and Spitz, 1977; Borys and Spitz, 1978). In addition several studies in the area of reading have also noted this difficulty with temporal relations (Shotick, 1960; Bos, 1979).

REFERENCES

Bennett, A. *A Comparative Study of Subnormal Children in the Elementary Grades*. New York: Teachers College, Columbia University, Bureau of Publications, 1932.

Blake, K. Aaron, I., & Westbrook, H. Learning of basal reading skills by mentally handicapped and non-mentally handicapped children. *Journal of Research and Development in Education*, 1969, *2*, 1-138.

Bliesmer, E. Reading abilities of bright and dull children of comparable mental ages. *Journal of Educational Psychology*, 1954, *45*, 321-331.

Borys, S. & Spitz, H. Reflection-impulsivity in retarded adolescents and nonretarded children of equal MA. *American Journal of Mental Deficiency*, 1978, *82*, 601-604.

Bos, C. *Inferential Operations in the Reading Comprehension of Educable Mentally Retarded and Average Students*. Unpublished doctoral dissertation, University of Arizona, 1979.

Byrnes, M. & Spitz, H. Performance of retarded adolescents and nonretarded children on the Tower of Hanoi problem. *American Journal of Mental Deficiency*, 1977, *81*, 561-569.

Cegelka, P. & Cegelka, W. A review of research: Reading and the educable mentally handicapped. *Exceptional Children*, 1970, *37*, 187-200.

Eamon, D. Selection and recall of topical information in prose by better and poorer readers. *Reading Research Quarterly*, 1978-79, *14*, 244-257.

Frederiksen, C. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458.

Frederiksen, C. Discourse comprehension and early reading. In L. Resnick & P. Weaver (Eds.) *Theory and Practice of Early Reading* (1) Vol. 1. Hillsdale, N.J.: Erlbaum, 1979.

Kintsch, W. *The Representation of Meaning in Memory*. Hillsdale, N.J.: Erlbaum, 1974.

Kirk, S., Kliebhan, J., and Lerner, J. *Teaching Reading to Slow and Disabled Learners*. Boston: Houghton Mifflin Co., 1978.

Madden, R., Gardner, E., Rudman, H., Karlsen, B. & Martin, J. *Stanford Achievement Test, Primary Level II* (NCS Edition). New York: Harcourt Brace Jovanovich, Inc., 1972.

Meyer, B. *The organization of prose and its effect on recall.* Amsterdam: North Holland Publishing Co., 1975.

Meyer, B. The structure of prose: Effects on learning and memory and implications for educational practice. In R. Anderson, R. Spiro, and W. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Erlbaum, 1977.

Meyer, B., Brandt, D., & Bluth, G. Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly,* 1980, *16,* 72-103.

Sheperd, G. Selected factors in the reading ability of educable mentally retarded boys. *American Journal of Mental Deficiency,* 1967, 71, 563-570.

Shotick, A. *A comparative investigation of the performance of mentally retarded and intellectually normal boys on selected reading comprehension and performance tasks.* Unpublished doctoral dissertation. Syracuse, New York: Syracuse University, 1970.

Spitz, H. & Borys, S. Performance of retarded adolescents and nonretarded children on one and two-bit logical problems. *Journal of Experimental Child Psychology,* 1977, *23,* 415-429.

Taylor, B. Children's memory for expository text after reading. *Reading Research Quarterly,* 1980, *15,* 399-411.

LINDA GAMBRELL

PATRICIA S. KOSKINEN
University of Maryland

JACK N. COLE
Prince George's County
Public Schools, Maryland

THE EFFECTS OF INDUCED MENTAL IMAGERY UPON COMPREHENSION: A COMPARISON OF WRITTEN VERSUS ORAL PRESENTATION

Induced mental imagery has been identified in recent research as a promising strategy for increasing comprehension of discourse (Kulhauy & Swenson, 1975; Lesgold, McCormick & Golinkoff, 1975; Levin, 1973; Steingart & Glock, 1979). Since mental imagery appears to be associated with efficient learning and remembering, it may provide a critical link in explaining the transition from a novice to sophisticated comprehender (Pressley, 1976, 1977). One cause for caution, however, with respect to interpreting the results of the research on mental imagery is that increases in learning from written text are slight while the results of learning from orally presented text have been more substantial. It appears that mental imagery has merit as a strategy for increasing listening comprehension but the value of induced mental imagery as a strategy for increasing reading comprehension needs further investigation (Tierney & Cunningham, 1980). The research to date on induced mental imagery does not provide a direct comparison of the effects of induced mental imagery upon listening and reading comprehension.

The main purpose of this study was to investigate the effects of induced mental imagery upon the comprehension of oral versus written discourse for above and below average readers. A secondary purpose was to determine, through a follow-up interview, the students' ability to induce mental imagery.

METHOD

Subjects

Sixty-three sixth grade above and below average readers enrolled in six suburban Maryland public schools served as subjects in the study. Criteria for inclusion in the study were: (1) reading comprehension scores obtained from the Iowa Test of Basic Skills (below average reading scores between 3.5 and 5.5 grade level and above reading scores of 6.5 +), (2) Cognitive Abilities Test scores within 1 standard deviation above or below the mean, and (3) teacher verification of subject identification as an above or below average reader.

Materials

The stimulus materials used in this study consisted of two short (approximately 300 words each) expository type passages written at the 3.0 grade level as determined by the Fry (1977) readability formula. A matrix design was used to develop parallel fictional passages about two groups of people. One passage was about the Blue Swamp Clan and a second passage was about the Pine Folks. The passages were developed to reflect the structure of materials which students encounter in the classroom; yet they contained novel information, thereby reducing the possible effects of prior knowledge.

For each passage a set of 10 short answer cued recall questions was constructed. Literal and paraphrase questions were asked for each passage ("What kind of work do the clan members do?", "Why did the Blue Swamp Clan leave their old home?").

A brief interview was developed for use following the listening and reading tasks. The purpose of the interview was to validate if students were able to induce mental imagery when given instructions to do so.

Procedure

Subjects were randomly assigned by ability to one of two conditions, listen-read and read-listen, depending upon whether they read or listened to a passage first. Half the students under each condition read the Blue Swamp passage and listened to the Pine Folks, while the reverse was true for the remaining subjects. These procedures were used to control for the effects of condition order and passage order.

The students were met individually for approximately 25 minutes. All subjects were instructed to "make pictures in your head to help you remember" under both listening and reading conditions. Immediately after listening to or reading a passage, the subjects responded to the interview about use of mental imagery. The interview did not tap specific passage information and therefore also served as an intervening activity to eliminate the effects of short term memory upon the comprehension assessments which followed. Upon completion of the interview, students were asked to retell the passage information. Subjects were told to pretend that they were trying to tell a friend everything they could remember about the Blue Swamp Clan (or Pine Folks). Subjects also responded to a 10-item cued recall test for each passage.

RESULTS

Two independent raters divided the two stimulus passages into propositions of three levels of importance using an adaptation of Meyer's (1975) procedure (interrater reliability for the Blue Swamp passage = .97, Pine Folks = .98). The subjects' free recall protocols for the listening and reading tasks were then scored by two independent raters (interrater reliability = .97).

Preliminary inspection of the data on free and cued recall revealed no differences attributable to story or order of presentation, so these variables were not considered further. Data on free and cued recall were analyzed using analysis of variance with repeated measures.

Significant ability differences were found, as expected, for both listening and reading tasks in favor of the above average readers on both free and cued recall. No statistically significant differences were found between the effects of induced mental imagery for listening and reading for either free (Table 1) or cued (Table 2) recall. On free recall of higher level propositions, however, sex and ability interaction effects were found for listening and reading. This sex and ability interaction effect was also found on the reading task for recall of lower level propositions (see Table 1).

TABLE 1

Means and Standard Deviations of Free Recall Scores for Propositions—Level 1 (highest), Level 2, and Level 3 (lowest)

		Males		Females	
		N=19 Above Average	N=15 Below Average	N=15 Above Average	N=14 Below Average
Level 1 Propositions* **					
Listening	M	10.32	6.73	9.67	5.79
	SD	3.63	2.91	3.41	4.22
Reading	M	10.00	7.07	8.47	7.71
	SD	3.48	4.68	2.99	2.86
Level 2 Propositions**					
Listening	M	7.16	5.00	6.87	5.86
	SD	3.33	2.64	3.72	3.65
Reading	M	7.95	6.00	6.53	5.29
	SD	4.07	4.15	3.13	2.75
Level 3 Propositions** ***					
Listening	M	1.47	1.00	.73	.50
	SD	1.12	.75	.70	.75
Reading	M	1.53	.53	.93	.79
	SD	1.30	.74	.96	.89

*significant ability differences at the .01 level.
**significant sex and ability interaction at the .05 level.
***significant ability differences at the .05 level.

TABLE 2

Means and Standard Deviations of Number of Correct Answers to Cued Recall Questions

		Males		Females	
		N=19 Above Average	N=15 Below Average	N=15 Above Average	N=14 Below Average
Listening*					
	M	9.05	6.66	8.00	7.85
	SD	2.39	2.41	2.17	2.14
Reading*					
	M	8.94	6.53	7.26	6.21
	SD	2.48	2.29	2.73	2.86

*significant ability differences at the .01 level.

DISCUSSION

The results of the comparison of the effects of induced mental imagery under listening and reading conditions are not consistent with the research which suggests that induced mental imagery is more potent for listening than for reading comprehension (Pressley, 1977). The results of the present study suggest that instructions to induce mental imagery may be equally effective under listening and reading conditions. It should be noted that this study attempted to eliminate and reduce major methodological problems inherent in the existing research which has prohibited direct comparisons of the effects of induced mental imagery for listening and reading comprehension of discourse.

Above average males recalled more higher level propositions than above average females, however, below average females recalled more than below average males under both listening and reading conditions. The same effect was found with respect to lower level propositions under the reading condition. The statistically significant sex and ability interactions found in this study suggest that sex differences may be a promising area for future research.

The results of the interview on use of mental imagery revealed that under the listening condition 17% of the below average subjects reported that they were not able to induce mental imagery while only 3% of the above average readers reported they were unable to do so. Under the reading condition 21% of the below average readers and 6% of the above average readers reported they were unable to induce mental imagery. More below average readers than above average readers report an inability to induce mental imagery. As Tierney and Cunningham (1980) have suggested, further research is needed to determine how imagery may be effectively induced since a rather large percentage of below average readers in this study apparently had difficulty imaging.

Finally, the results of this investigation lead to two additional conclusions. First, the results support the findings of Smiley et al. (1977) and Guthrie (1973) which suggest that poor readers suffer from specific comprehension difficulties which do not appear to involve decoding skills. There were large differences between above and below average readers comprehension performance under *both* listening and reading conditions. Second, the data are also consistent with the assumption that listening and reading comprehension depend upon the same basic process.

ACKNOWLEDGMENTS

The research reported herein was supported in part by a grant to the first author from the University of Maryland Reading Center Research Committee.

The authors would like to thank Natalie Felsher and Paula Jawitz for their help in data collection and Dr. Chan Dayton and Gary Skaggs for statistical assistance.

REFERENCES

Fry, E. Fry's readability graph: Clarification, validity, and extension to level 17. *Journal of Reading*, 1977, *21*, 242-252.

Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. *Journal of Educational Psychology*, 1973, *65*, 294-299.

Kulhauy, R. W., & Swenson, I. Imagery instructions and the comprehension of text. *British Journal of Educational Psychology*, 1975, *45*, 47-51.

Lesgold, A. M., McCormick, C., & Golinkoff, R. M. Imagery training and children's prose learning. *Journal of Educational Psychology*, 1975, *67*, 663-667.

Levin, J. R. Inducing comprehension in poor readers: A test of a recent model. *Journal of Educational Psychology*, 1973, *65*, 19-24.

Meyer, B. J. F. *The organization of prose and its effect on memory*. Amsterdam: North Holland, 1975.

Pressley, G. M. Mental imagery helps eight-year olds remember what they read. *Journal of Educational Psychology*, 1976, *68*, 355-359.

Pressley, M. Imagery and children's learning: Putting the picture in developmental perspective. *Review of Educational Research*, 1977, *47*, 585-622.

Smiley, S. S., Oakley, D. D., Warthen, D., Camprone, & Brown, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Journal of Educational Psychology*, 1977, *69*, 381-387.

Steingart, S. R., & Glock, M. D. Imagery and the recall of connected discourse. *Reading Research Quarterly*, 1979, *15*, 66-83.

Tierney, R. J., and Cunningham, J. W. Research on teaching reading. In P. D. Pearson (Ed.), *Handbook on research in reading*. Longman, in press.

CONNIE A. BRIDGE
SUSAN THOMAS
LINDA EVERS
University of Kentucky

THE EFFECTS OF TEXTUAL COHESION ON CLOZE COMPREHENSION SCORES

The purpose of this study was to examine the effect of cohesion upon subjects' ability to fill in deleted words in a cloze passage. In a prevous study, Thomas & Bridge (1980) had found a highly significant correlation (.80) between subjects' cloze scores and their ability to employ macrostructure operations in the generation of summaries. A detailed analysis of the number and type of propositions in subjects' summaries indicated that propositions which were most frequently recalled contained numerous cohesive ties (as defined by Halliday & Hasan, 1976) with preceding idea units and served either an integrative or generalization function (as defined by van Dijk, 1977 and Kintsch, 1976) for the related propositions in the text. Thus, it was suggested that since the analysis of cohesive ties can help explain the recall of certain propositions in a text perhaps an analysis of these ties may also be able to help explain why certain words in a cloze passage are or are not likely to be filled in correctly.

ANALYSIS OF COHESION

To investigate this question it was necessary to delete from the cloze passages target words which were involved in the cohesive relationships within the text. The first step in this process involved an analysis of the cohesion of the text employing the procedures described by Halliday and Hasan (1976). The passage used in the study was a 572 word excerpt from an expository passage by S. E. Brock entitled "The Cat that Kills with a Single Spring" taken from the *Reader's Digest New Advanced Reading Skill Builder I*, p. 78-82. According to the Fry Readability Graph, the readability level was low seventh grade. The structure of the text was of the descriptive listing type in that it simply described various characteristics and habits of the jaguar. Most of the content was judged to be unknown to the majority of readers without having read the passage. Using the Halliday and Hasan coding scheme, the cohesive ties were identified. Each cohesive item was listed along with the presupposed item to which it was related. The type of cohesion (reference, substitution, ellipsis, conjunction, and lexical) was noted as well as the distance, which was found by counting the number of intervening sentences between the cohesive item and its presupposed item.

This analysis of cohesion revealed no cohesive items which could be classified as elliptical and only one tie that involved substitution. Ellipsis and substitution tend to occur more often in dialogue, and there was no dialogue in this passage. There were ten ties that involved conjunction, 26 referential ties, and 112 lexically related ties identified through the use of the coding scheme. Of course, all texts which are about a specific topic will tend to have a high percentage of lexical items which perform a cohesive function. Sticking to a topic dictates the occurrence of reiteration of the same word, a synonym, a superordinate, or a general word referring back to a presupposed item and of collocation which results from the co-occurrence of words that are typically associated with each other because they tend to appear in similar contexts. Because this text, as previously mentioned, was a series of paragraphs about jaguars, there were many lexical items of the reiterative and collocative type referring to "jaguar", "big cats", and "killing."

Deletion Patterns

After the analysis of cohesion was completed, the second step involved selecting the target items to be deleted. An attempt was made to identify pairs of cohesively related words; that is, a cohesive item and its presupposed item would be considered a cohesive pair. Because of the richness of cohesion in this passage, approximately thirty-six percent of the words had been identified as either a cohesive item, a presupposed item for one or more cohesive items, or both a cohesive item and a presupposed item. In addition to this thirty-six percent were the words in the whole sentences which were related by the conjunctive elements. Because whole sentences were related by the conjunctive items, it was not possible to specify simply a pair of words as a cohesive and presupposed item-pair. Thus, a decision was made to look only at the lexically and referentially related item pairs in this experiment and to study them separately to avoid the confounding effect of having an item that was part of both a referential and lexical cohesive pair.

For the referentially-related item pairs, 16 pairs of words were identified. Two versions of the passage were developed. In one version, both the cohesive item and its presupposed item were deleted. This was labeled the Reference-Related deletion pattern. In the second version, the cohesive item was deleted but the presupposed item was left intact, and the word immediately following the presupposed item was deleted. This was known as the Reference-Unrelated deletion pattern. If cohesion affects readers' comprehension as predicted, then readers should have more difficulty filling in the deleted words when both words in the cohesive pair are deleted. Furthermore, in the related condition, the subjects' ability to fill in the cohesive item, the second item of the pair, should also be affected by whether or not they were able to successfully supply the first or presupposed word. Since twelve of the sixteen pairs involved a pronoun and its antecedent, there was a strong expectation that the ability to supply the presupposed item or antecedent would influence subjects' ability to supply the cohesive item or pronoun. If so, this effect would be exaggerated by the fact that three of the pronouns had the same presupposed item in one paragraph, while four pronouns referred to the same presupposed item in another paragraph. A total of twenty-seven words was deleted from each of the passages designed to test the effect of referential cohesion on cloze comprehension scores.

Two more versions of the same passage were developed to study the effect of deleting lexically related pairs of words upon subjects' ability to fill in cloze deletions. Fourteen word pairs, totaling twenty-eight words, were deleted. In the Lexical-Related version, both the lexical cohesive item and its presupposed item were deleted. In the Lexical-Unrelated version, the lexical cohesive item was deleted, but the presupposed item was left intact and the word immediately following it was deleted. Again, it was predicted that the deletion of the related presupposed items should disrupt subjects' comprehension more than the deletion of words that were not involved in the cohesive relationship and that the ability to supply the presupposed item in the related version should affect the subjects' ability to supply the target cohesive item.

Subjects

Sixty-eight ninth graders enrolled in language arts classes in a small West Virginia city were randomly assigned to read and complete one of the four versions of the cloze passage. Their responses were scored for exact replacement and semantic acceptability.

Results

An analysis of subjects' scores on the four versions of the cloze passages did not fully support the hypothesis that subjects' cloze comprehension scores would be adversely affected by the deletion of the presupposed item in the lexically and referentially related pairs.

Table 1
Percentage Scores on the Related and
Unrelated Conditions of the Reference and Lexical Deletions

		Lexical		**Reference**	
		Exact	**Semantic**	**Exact**	**Semantic**
Related	$\bar{X}$	.32	.54	.43*	.76
	S.D.	.11	.16	.14	.18
Unrelated	$\bar{X}$	.43**	.60	.34	.68
	S.D.	.13	.19	.12	.15

* $P < .05$
** $p < .01$

The differences were in the expected direction in the passages containing lexical deletions in that subjects correctly supplied 32% of the missing items in the related deletion pattern and 43% in the unrelated pattern when their responses were scored for exact replacement. This difference was significant at the .01 level. When given credit for semantically acceptable responses their scores were 54% on the related and 60% on the unrelated. However, this difference was not significant.

The results on the passage with the reference items deleted ran counter to predictions, in that subjects scored significantly ($p < .05$) higher (43%) in the related deletion pattern than they did in the unrelated deletion pattern (34%) when responses were scored for exact replacement only. Using the semantically acceptable criteria, their scores were 76% on the related and 68% on the unrelated deletion. This difference was not significant. They actually scored higher when the presupposed item of the cohesive pair was deleted, but the difference was only significant when scored for exact replacement.

In an attempt to explain these unexpected results, further analysis of the words affected by the two reference deletion patterns was conducted. Most of the referential pairs involved pronominal reference with many of the pronouns referring back to either the singular or plural form of jaguar. Because of the frequency with which jaguar was mentioned in the passage, subjects were highly successful in using the redundancy of the passage to correctly supply the term "jaguar" or "jaguars." In the exact replacement condition 9 of the 17 subjects correctly supplied "jaguars" which was the presupposed item for 5 of the involved referential pairs and 13 of 17 subjects supplied "jaguar", the presupposed item for 3 other referential pairs. It was postulated that in the unrelated referential condition the deleted words appeared less frequently in the passage and were less predictable from the context. However, separate comparisons of subjects' scores on the presupposed reference items and the cohesive items did not support this explanation.

Table 2
Percentage of Presuppose Items
Correct in Related & Unrelated Conditions

		Lexical		References	
		Exact	**Semantic**	**Exact**	**Semantic**
Related	$\bar{X}$	.28	.49	.50	.74
	Percentage of Presuppose Correct in Related & Unrelated Conditions	1.41*	.62	.46	.73
	S.D.	.17	.22	.18	.19

*p < .05

Table 3

Percentage of Cohesive Items Correct
In Related & Unrelated Conditions

		Lexical		Reference	
		Exact	**Semantic**	**Exact**	**Semantic**
Related	$\bar{X}$	.39	.61	.41**	.79
	S.D.	.14	.19	.14	.20
Unrelated	$\bar{X}$	.43	.56	.27	.70
	S.D.	[illegible]	.16	.11	.14

**p < .01

When scored for exact replacement, the subjects scored slightly lower on the less frequent cohesively unrelated words, but the difference did not approach significance. When scored for semantic acceptability, their scores were almost identical. A comparison of subjects' scores on the cohesive items which involved the same words in both conditions showed that subjects scored less well on these items in the cohesively unrelated condition than in the related condition even though there were no differences in their scores on the presupposed items in the related and unrelated conditions. Thus, it appears that factors other than the cohesive relationship between the presupposed word and its cohesive word affected subjects' ability to supply the cohesive words.

A further analysis of the word pairs in the related and unrelated conditions was conducted to examine the relationship between subjects' ability to correctly supply the first word in the pair and their ability to supply the second. For each pair of cohesive items it was possible for one of four combinations to occur: 1) both items could be right; 2) the first item could be right, the second wrong; 3) the first could be wrong, the second right; or 4) both could be wrong. The following table summarizes this data by presenting the relative frequencies with which each of these events occurred in the related and unrelated conditions scored for semantic acceptability and exact replacement.

Table 4

Relative Proportions of Right/Wrong
Combinations in Related & Unrelated Item Pairs

	Reference Pairs			
	Semantic		Exact	
	R-R	R-U	R-R	R-U
Both right	.66	.54	.22	.18
First right/second wrong	.08	.18	.28	.28
First wrong/second right	.13	.16	.18	.09
Both wrong	.13	.12	.32	.45
$\chi^2_3 = 13.46$ $p < .005$			$\chi^2_3 = 14.75$ $p < .004$	
	Lexical Pairs			
	Semantic		Exact	
	L-R	L-U	L-R	L-U
Both right	.38	.37	.15	.18
First right/second wrong	.11	.24	.13	.22
First Wrong/second right	.23	.19	.23	.25
Both wrong	.28	.20	.45	.35
$\chi^2_3 = 15.83$ $p < .003$			$\chi^2_3 = 12.84$ $p < .007$	

For the cohesively related pairs in the related conditions, only combination three would be inconsistent with the notion that the presupposed item must be correctly supplied before the cohesive item could be supplied. This occurred in 23 percent of the item pairs in the Lexical-Related condition under both methods of scoring. In the Reference-Related condition this combination occurred in 13 percent of the item pairs when the semantically acceptable scoring method was used and 18 percent of the pairs when scored for exact replacement. Thus, in 77% of the lexical pairs and in more than 82% of the reference pairs, the subject was unable to supply the cohesive item if the presupposed item had not been correctly supplied.

If the cohesive relationship between the words in the item pairs was the primary factor influencing the right/wrong combinations, then the item pairs in the unrelated condition should not be affected in the same way. There would not be any reason to expect fewer cases of combination three, first item wrong/second right than of combination two, first item right/second wrong, because the student's ability to get the second item should not be directly related to their success in supplying the first item. Chi square tests comparing the related and unrelated distributions revealed that the distributions were significantly different from each other. However, these differences did not appear to be the result of the cohesive relationships involved in the related condition, as in two of the comparisons there were fewer instances of combination three in the unrelated condition than in the related condition and in the other two comparisons almost the same proportion of instances of combination three occurred in both the related and unrelated conditions. Thus, it appeared that factors other than the referential or lexical cohesive relationship was influencing subjects' ability to successfully supply deleted cloze items.

CONCLUSIONS

In this experiment, different versions of an expository passage were created in which the various cloze deletion patterns were based upon the identification of lexically or referentially related words pairs. Subjects' scores on these passages did not reveal a systematic effect resulting from the lexical or referential cohesive relationships identified in the analysis of cohesion. Thus, it was concluded that factors other than the cohesive relationships were influencing subjects' ability to supply the cloze deletions.

IMPLICATIONS

In his newly released and updated review of cloze instruction research, Jongsma (1980) suggests that researchers may be able to use Halliday & Hasan's cohesion system as a guide in selectively deleting the words in cloze instruction materials. By focusing on certain types of cohesive ties, Jongsma suggests that it might be possible to improve students' ability to recognize and comprehend the targeted semantic relationship and thus ultimately improve their comprehension. He points out that this approach could be easily adapted to expository materials in content areas.

Although the results of this experiment indicate that the subjects in the study were not consistenlty using the cohesive relationships in the target passage, it might be worthwhile to design other cloze passages which focus on specific cohesive relationships using deletion patterns similar to the ones employed in this study and to provide instruction to students which enables them to consciously practice identifying and using cohesive relationships. Further research will be required to investigate the feasibility of using cohesion analysis as a framework for selectively deleting items in cloze instruction materials.

REFERENCES

Halliday, M. A. K., & Hasan, R. *Cohesion in English.* London: Longman Group Limited, 1976.
Jongsma, E. A. *Cloze instruction research: a second look.* Newark, Delaware: International Reading Association, 1980.

Kintsch, W. Comprehension and memory of text. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 5). Hillsdale N. J.: Lawrence Erlbaum Associates, 1978.

Thomas, S., & Bridge, C. A comparison of subjects' cloze scores and their ability to employ macrostructure operations in the generation of summaries. In M. L. Kamil & A. J. Moe (Eds.), *Perspectives on reading research and instruction*. Twenty-Ninth Yearbook of the National Reading Conference. Washington, D.C.: National Reading Conference, 1980, 69-77.

van Dijk, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

JILL FITZGERALD WHALEY
University of North Carolina
at Chapel Hill

READERS' REACTIONS TO TEMPORAL DISRUPTION IN STORIES

Recent research has explored the question of whether individuals have a concept of idealized story structure; i.e., whether readers and listeners have a mental representation of story elements, such as beginning, development of action, and ending, and of how these elements fit together (e.g., Mandler & Johnson, 1977; Rumelhart, 1978; Stein & Glenn, 1979; Thorndyke, 1977). Findings from investigations based predominantly (though not entirely) on recall tasks have been very consistent in demonstrating that 1) individuals do appear to have an idealized concept of story, 2) individuals' knowledge of story structure develops over time, and 3) readers and listeners appear to use this knowledge about story parts to understand and remember stories (cf., Chodos & Mosenthal, 1978; Mandler, 1978; Mandler & Johnson, 1977; Rumelhart, 1978; Stein, 1976, 1979; Stein & Glenn, 1979; Thorndyke, 1977; Whaley, 1979). However, there has been occasional contradictory evidence. At least two studies, by Baker (1978) and by Stein and Nezworski (1978), indicate that different research tasks may vary in the extent to which they reveal individuals' knowledge of story structure.

The purpose of the present study was to further explore individuals' concept of story structure, but to do so from yet another perspective, i.e., using procedures other than recall tasks. The rationale for design of the tasks was that if readers *expect* a particular ideal structure to occur in narrative materials, then, when they are reading, they ought to be aware of temporal and structural disruption in stories, i.e., they should be startled by violations of those expectations. Also, more extreme violations should be more noticeable. Thus, two tasks were designed so that subjects read stories in which story parts had been moved around or deleted, and then the subjects indicated orally if something about the story structure disturbed them. These tasks are similar to ones used by Stein (1976) and Stein and Glenn (1977) except Stein and Glenn asked subjects to listen to stories having violations and then recall the stories. Their focus was on the effect of the violations on recall. In the present study, the focus was on readers' expression of their awareness of disruptions in stories.

Three main questions were addressed in the present investigation: 1) Do readers express an awareness of structural disruption in stories? 2) Are there developmental differences in the extent to which readers are aware of structural disruption? and 3) Are there developmental differences in the types of structural disruption that disturb readers?

METHOD

Subjects

Fifty third graders, 51 sixth graders, and 52 eleventh graders completed Task I. The same sixth and eleventh graders participated in Task II. Only students reading at or above grade level were selected for the study. This was done at the lower grade levels to control for possible word recognition difficulties and to insure at least literal level comprehension. Then, in order to make comparisons across grades, it seemed reasonable to maintain consistent rules for reading ability at each grade level. Reading level was determined using teacher judgement and a previously administered standardized test. Different tests were used at different grade levels (*Gates-MacGinitie Reading Tests*, Gates & MacGinitie, 1965; *Iowa Test of Basic Skills*, Lindquist, Hieronymus 1955-56; *Tests of Academic Progress*, Scannell & Smith, 1964-66).

Materials

The Task I story was a variant of "Fox and Bear" (cf., Stein and Glenn, 1979); the Task II story was a modified version of "Harry by the Sea" (Zion, 1965). Using the Spache (1974) formula, readability levels of the two stories were 2.6 and 2.5, respectively. The Mandler and Johnson (1977) story grammar was used to identify basic story parts in each story: Setting, Beginning (a precipitating event), Complex Reaction (the protagonist's reactions to the precipitating event and establishment of a goal, Attempt (the attempt to achieve the goal), Outcome (the success or failure of the attempt), and Ending. For identification of these parts in the stories, the interrater reliability was .89.

Procedures

In Task I, the Reaction and the Outcome were systematically moved forward or backward in the story. These two categories were selected because in prior research the Reaction category was consistently least often recalled and the Outcome was one category most frequently recalled. Thus, when moved, these two categories might represent ones that could be less noticeable and more noticeable, respectively. Subjects were randomly assigned to one of nine conditions: 1) the Reaction was moved toward the start of the story by one category so that it preceded the Beginning; 2) the Reaction was moved toward the start of the story by two categories; 3) the Reaction was moved toward the end of the story by one category so that it followed the Attempt; 4) the Reaction was moved two categories toward the end; 5) the Outcome was moved one category toward the start; 6) the Outcome was moved two categories toward the start; 7) the Outcome was moved one category toward the end; 8) the story parts were scrambled in random order; and 9) the story was presented intact (control condition).

In Task II, subjects were randomly assigned to one of seven conditions. In conditions one through six, the six story categories (Setting, Beginning, Reaction, Attempt, Outcome or Ending) were systematically deleted, one for each condition. Condition seven was a control situation in which the entire story was read. No spaces or blank lines were inserted in the material that the students read.

Students read the stories individually. They were told to read each story silently, and if they encountered something wrong with the way the story was put together, they were to tell the examiner about it. They were told that one or more sentences might have been moved around in the story or that one or more sentences might have been removed. They were also warned that the material may *not* have been tampered with. After completing the tasks, subjects were asked if they had heard either of the stories before. None of the students indicated they had heard them.

Analysis

To analyze subjects' responses, a score of two was given when the child indicated something was wrong with the structure of the story. The child had to at least identify the appropriate place of the category shift or deletion. In Task II, the subject did not have to supply the missing information. A score of one meant the student thought the structure of the story was acceptable. Thus, for Task I, each subject had a final score of one or two; and for Task II, each subject again had a final score of one or two. Interrater reliability for scoring protocols was .95 for Task I and .94 for Task II.

Two separate sets of two-way (grades by conditions) univariate analyses of variance were performed, one for each task.

RESULTS

To determine whether the readers expressed an awareness of the structural changes made in the stories, the main effects for conditions were ascertained for each task. In Task I, there was a significant main effect for conditions ($F = 7.76$; $p < .0001$). The means for conditions one through nine were 1.63, 1.89, 1.30, 1.22, 1.22, 1.43, 1.33, 1.73, and 1.07, respectively. However, planned contrasts comparing each condition to the control group indicated that readers were *not* disturbed by many of the category displacements. Specifically, only conditions one, two, six, and eight were significantly different from the control condition. The t statistics and probability levels for these planned contrasts were: 4.01, $p < .01$; 5.83, $p < .01$; 2.44, $p < .05$; 4.59, $p < .01$; respectively. In other words, movement of the Reaction category one and two nodes towards the start of the story and movement of the Outcome two nodes towards the start of the story were noticed. Also, nearly all students were aware of severe disruption in the scrambled story. In most cases, then, moving a category toward the start of the story *was* noticeable; moving it towards the end was not noticeable. Also, increasing the distance of category displacement significantly affected readers' awareness in only one instance. Moving the Reaction toward the start of the story by two categories was significantly more noticeable than moving it one category toward the start ($t = 2.00$, $p < .05$).

In Task II, there was also a significant main effect for conditions ($F = 9.06$, $p < .0001$). The means for conditions one through seven, respectively, were 1.37, 1.82, 1.07, 1.13, 1.15, 1.00, and 1.34. Tests of planned contrasts showed only the responses in condition two (where the Beginning was deleted) were significantly different from those in the control situation ($t = 4.80$, $p < .01$). In other words, on the average, except for the case of the deleted Beginning, readers were *not* cognizant of category deletions. This is not surprising in the case of the Reaction deletion because explicit statements of Reactions are considered optional in Mandler and Johnson's (1977) grammar. The result *is* surprising in the other cases.

To determine whether there were developmental differences in the extent to which readers were aware of structural disruption, the main effects for grades were ascertained for each task. In both tasks there were significant main effects for grades (Task I, $F = 16.77$, $p < .0001$; Task II, $F = 8.14$, $p < .0055$). In Task I, eleventh graders noticed the category shifts significantly more often than did third graders ($t = -5.18$, $p < .001$), and sixth graders were significantly more aware of the shifts than were third graders ($t = -4.20, p < .001$). There was no difference between sixth and eleventh graders ($t = .97, p > .05$). In Task II, eleventh graders noticed category deletions significantly more often than did the sixth graders ($t = -3.13, p < .01$). In general then, for both tasks, older students were more aware of temporal disruption than were younger students.

Finally, the interactions of grades by conditions were analyzed to determine if there were developmental differences in the types of structural disruptions that might disturb readers. The interactions were not significant in either task (Task I, $F = 1.25, p < .2431$; Task II, $F = .65, p < .6928$).

CONCLUSIONS AND DISCUSSION

The main conclusions drawn from the results of the two tasks used in this study were: 1) the readers did not always express awarness of structural disruption, so it appeared that they did not always expect each of the story categories to be present, and they did not consistently expect Reaction and Outcome categories to appear in specific places; 2) older children were more aware of temporal and structural disruption than were younger children; and 3) there were no developmental differences in children's awareness of various types of structural variations.

The strength of the interpretation of the findings is tempered by the fact that only one story was used in each task. It is possible that the results may be story specific. Also, only two categories were manipulated in Task I. However, across the two tasks, the results do tend to point to the same conclusions.

With these comments in mind, there are two main points that might be discussed. First, because the results are somewhat at odds with previous research, these findings may indicate a "task effect;" these tasks may not be good measures of readers' knowledge of story structures *or* they may tap a true, but separate dimension of story structure knowledge.

Secondly, the results might be compared to the two previous studies by Stein (1976) and Stein and Glenn (1977) in which categories were moved or deleted and the effect on recall was analyzed. The conditions under which accuracy of *recall* was affected in Stein (1976) and Stein and Glenn's (1977) studies do *not* correspond exactly to the conditions in the present study under which children were *aware* of structural variation. To give one example, movement of the Reaction category so that it preceded the Beginning *was noticed* in the present study, but under the same category movement in Stein's (1976) study, accuracy of recall was *not* decreased. Taken together, one interpretation of the results of these three studies is that metacognition regarding story structures may be unrelated to comprehension and accurate recall. This is an issue that specifically deserves to be addressed in further studies.

REFERENCES

Baker, L. Processing temporal relationships in simple stories: Effects of input sequences. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 559-572.

Chodos, L. & Mosenthal, P. Fourth graders' comprehension of story structures under three recall conditions. In P. D. Pearson, and J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice*, 27th Yearbook of the National Reading Conference, 1978.

Gates, A. I., & MacGinitie, W. H. *Gates-MacGinitie Reading Tests*, N. Y.: Teachers College Press, 1965.

Lindquist, E. F., & Hieronymus, A. N. *Iowa Test of Basic Skills*. Boston: Houghton Mifflin, 1955-56.

Mandler, J. M. A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1978, *1*, 14-35.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

Mosenthal, P. Fourth graders' comprehension of story structures under three recall conditions. In P. D. Pearson, and J. Hansen (Eds.), *Reading: Disciplined inquiry in process and practice*, 27th Yearbook of the National Reading Conference, 1978.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge and J. Samuels (Eds.), *Basic procedures in reading: Perception and comprehension*. Hillsdale, N. J.: Lawrence Erlbaum, 1978.

Scannell, D. P. & Smith, H. P. *Tests of Academic Progress*. Boston: Houghton Mifflin, 1964-66.

Spache, G. *Good reading for poor readers*. Champaign, Ill.: Garrard, 1974.

Stein, N. L. How children understand stories: A developmental analysis. In L. Katz (Ed.), *Current topics in early childhood education* (Vol. 2). Hillsdale, N. J.: Ablex, 1979.

Stein, N. L. *The effects of increasing temporal disorganization on children's recall of stories*. Paper presented at the Psychonomic Society Meetings, St. Louis, 1976.

Stein, N. L., & Glenn, C. G., An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (Vol. II in series, Advances in Discourse Processes). Norwood, N. J.: Ablex, 1979.

Stein, N. L., & Glenn, C. G. *The role of structural variation in children's recall of simple stories*. Paper presented at the Society for Research in Child Development, New Orleans, 1977.

Stein, N. L., & Nezworski, G. The effects of organization and instructional set on story memory. *Discourse Processes*, 1978, *1*, 177-193.

Thorndyke, P. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 97-110.

Whaley, W. J. *An analysis of readers' expectations for story structures*. Doctoral dissertation, State University of New York at Buffalo, 1979.

Zion G. *Harry by the sea*. N. Y.: Harper & Row, 1965.

THOMAS W. BEAN
ASHLEY BISHOP
California State University

SELECTED PROSE CHARACTERISTICS OF AN ESL NARRATIVE AND THEIR EFFECT ON COMPREHENSION

Story comprehension is influenced by a complex array of psycholinguistic features. For example, readers possess a culturally determined schema for stories as problem-solving episodes that resemble real life experiences (Bower, 1978; Tannen, 1980). Numerous studies confirm that readers display a well-developed sense of story structure when they are asked to retell a narrative (Mandler and Johnson, 1977; Rumelhart, 1977; Dreher and Singer, 1980; Stein and Glenn, 1979).

This internal sense of story structure is derived from countless exposures to predictable, temporally sequenced narrative categories including: 1) setting; 2) initiating event; 3) development; 4) protagonist's goal formation; 5) attempt to attain the goal; and 6) a response to the attempt. The categories vary in structural importance as evidenced in readers' recall summaries. For example, the gist or goal of a story is structurally more important than the setting (Bower, 1978), though not necessarily more memorable (Stein and Glenn, 1979). A narrative may deviate from a reader's expectations at the story structure level or at the level of individual propositions within categories. When a narrative departs from a reader's expectations at either level, processing time is increased and recall inhibited (Mandler, 1978; Stein and Nezworski, 1978; Thorndyke, 1977). Similarly, trivial propositions within a narrative may be highly salient for some readers, thereby limiting retention of more important story features such as the gist (Bean, Potter, and Clark, 1980). This notion of propositional salience or "relevance assignment" (van Dijk, 1979, p. 113) is just beginning to receive attention in the analysis of children's story recall (cf. Stein and Glenn, 1979). It is particularly pertinent to the results of the study described in this paper.

In addition to overall story structure and propositional salience, a host of semantic features operating at the sentence and intersentential level may encourage or inhibit comprehension. For example, one semantic feature that has intrigued prose researchers in recent years is pronoun reference. Recent studies suggest that an author's use of pronouns comprises a tacit contract with the reader that assumes facility in making any bridging inferences required to identify appropriate antecedents (Nash-Webber, 1978; Warren, Nicholas, and Trabasso, 1980). Early sentence level studies concluded that pronoun ambiguity was problematic for middle grade students (Richek, 1977). However, more recent studies exploring pronoun resolution in continuous, well-formed narratives suggest that this feature of prose may be much less problematic than previously believed (Barnitz, 1980).

The majority of story structure and pronoun reference studies have involved standard English (SE) speaking subjects and story material atypical of basal reader selections. In contrast, Bean et al. (1980) explored the effect of systematically modifying story structure propositions and pronoun antecedents on bilingual (Spanish-English) students' comprehension of a basal selection. A story entitled "The Chase Twins" from the *Miami Linguistic Readers* (D.C. Heath, 1966) comprised the stimulus material. Adjusting the story to make 20 of the most obscure pronoun referents explicit failed to significantly improve either students free or probed recall scores. In contrast both probed and free recall scores revealed that students were able to comprehend the gist or goal of the story significantly better when trivial digressions were removed, resulting in a predictable, problem-solving scenario. Moreover, students' comprehension was significantly higher on the restructured version, despite its increased decoding difficulty as measured by the Raygor (1977) readability graph. This led the authors to conclude that the students profited from story material that encouraged comprehension, not just decoding ease.

The present study is a replication of the aforementioned research with a sample of monolingual (English) children. The findings of this replication depart from the previous study yet serve to reinforce earlier conclusions with respect to the bilingual students' comprehension of the story. Present findings will be discussed in light of salience of particular propositions within story categories and free versus probed recall measures.

METHOD

Subjects

Thirty fourth graders attending a public elementary school in a middle-class suburb of Orange County, California were involved in the study. All subjects were average readers based on teacher judgment and placement in stanines four, five or six of the *Nelson Reading Test* (Houghton Mifflin, 1977).

Materials

The opening portion of "The Chase Twins" (D.C. Heath, 1966) had been parsed into propositions and story structure categories in the previous study. This section contained the first four categories of Stein and Glenn's (1979) story grammar: 1) setting; 2) initiating event; 3) development; and 4) goal formation. The gist or goal of the story centered on twins who planned to play detective and find a stolen ruby ring. However, the predictable movement of the story was continually disrupted by the introduction of trivial propositions. For example, in the setting category, the twins argue about what to play that day—cops and robbers or hospital. This seemingly trivial proposition formed a highly salient sub-goal for the children in the previous study. Their recall protocols indicated that the majority of their attentional resources were devoted to the least important aspects of the story in terms of its gist (Bean, Note 1).

The present study employed the three story conditions explored in the previous research. Table 1 provides an overview of the length, readability, and

TABLE 1

Story condition, length, readability, and propositional density by category

				Propositions Per Category			
Story Condition	# of Words	Readability	Propositions	Setting	Init Event	Dev	Goal Formation
Original	583	3.0	98	13	8	56	21
Pronoun Rewrite	624	4.0	100	15	8	56	21
Story Rewrite	427	5.0	61	2	7	40	12

propositional density by category for each of the three story conditions. The readability of each story condition was determined with the Raygor (1977) graph. Thorndyke's (1977, p. 82) definition of a proposition as "a clause or sentence containing an action or stative verb" was used to guide the parsing of the story into propositions.

The original version of "The Chase Twins" comprised the first condition. The second condition was rewritten at the sentence level to provide explicit antecedents for 20 obscure pronouns without significantly altering the propositional density of the original categories. For example, the development category of the original story contained the single proposition and related pronoun "*it* really is a twin mystery." This was rewritten as two propositions: *"the stolen ruby ring* is a twin mystery." These sentence level pronoun edits resulted in a slightly longer version than the original story and increased the readability by one grade level.

The third story condition involved removing trivial propositions that seemed to detract from a predictable problem-solving structure. For example, the original story setting contained 13 propositions. These were deleted leaving just two introductory propositions in the rewritten version. Pronoun referents were made explicit as in the second story condition. These modifications raised the readability two grade levels beyond that of the original.

Procedure

The 30 students were randomly assigned to one of the three story conditions resulting in 10 subjects per condition. They met individually with the researchers and were instructed to read the story silently so they could retell it in their own words. Their retellings were taped and followed by a 10 item probed recall containing wh-questions.

The probed recall was designed to assess subjects' comprehension of the gist or goal formation category (i.e. find the missing ring and solve the mystery), as well as their success in identifying obscure pronoun referents. For example, question nine, "What do the twins decide to do?" pertained to the goal formation category. Question six, "Why does Mr. Chase know so much about the robbery?" (he is one of the policemen who caught the robber) revealed the degree to which subjects were able to resolve pronoun referents. Both probed and free recalls were scored by two independent raters yielding inter-rater reliability coefficients of $r > .90$.

Subjects' free recalls were transcribed and compared to the propositional analysis for each story condition. Free recalls were scored liberally for the number of propositions produced verbatim or in paraphrase form in each story category. Free recall scores for each subject were then transformed into proportions based on a ratio of the subject's total propositional recall to his/her recall in each category. This provided an indication of propositional salience in each category and story condition. Table 2 presents the mean proportion of propositions recalled in each story condition and category.

RESULTS

The mean probed recall scores for each of the three story conditions were: 3.47 on the original (s.d. = 1.36); 4.70 on the pronoun rewrite (s.d. = 1.50); 5.62 on the story rewrite (s.d. = 1.64). An analysis of variance indicated a significant

TABLE 2
Mean proportion of recall by story condition and category

Story Condition	Category			
	Setting	Initiating Event	Development	Goal Formation
Original	.46	.06	.22	.26
Pronoun Rewrite	.16	.10	.23	.38
Story Rewrite	.02	.18	.46	.31

difference in probed recall scores across the three story conditions (F (2,27) = 5.11, $p < .025$). Scneffé comparisons of the group means revealed a significant difference for the story rewrite condition and the original version (difference = 2.15, $p < .025$). Comparison of the pronoun rewrite condition and the original version did not reach conventional significance (difference = 1.23), although it did approach significance at .10. As in the previous study, providing explicit pronoun referents resulted in a negligible contribution to students' comprehension. While these results seemed to parallel the probed recall data from the previous study, free recall protocols provided a more detailed, and perhaps more accurate picture of this group of subjects' narrative processing, particularly in terms of story gist.

An analysis of variance comparing subjects' free recall of the goal formation category was conducted for the three story conditions. In contrast to the bilingual subjects who achieved their greatest gist recall in the story rewrite condition, there was no significant difference in recall of goal propositions for these students regardless of story condition. Table 2 confirms the relative balance of goal recall across the three conditions for these subjects. The remainder of the paper will be devoted to a discussion of these contrasting, cross-sample results with a focus on propositional salience and the task demands of free versus probed recall.

DISCUSSION

In the previous study (Bean et al., 1980), bilingual subjects allocated their attentional resources to the sibling argument in the setting category of the original story version (Figure 1). This salient event formed a sub-goal that detracted from their comprehension of the gist. Moreover, bilingual students' probed and free recall results (Bean, Note 1) confirmed the powerful effect of deleting trivial events on their comprehension.

While the probed recall data suggest that subjects in the present study also benefit from a predictable story structure, the free recall protocols show that these subjects deployed their attentional resources in a more diffuse fashion. For example, in Figure 1, it can be seen that both groups of subjects recalled a substantial number of the 13 setting propositions. In contrast, their recall of the protagonists' goal differs dramatically. In the story rewrite condition, bilingual subjects' recall of the protagonists' goal parallels that of the present subjects (Figure 2). Conversely, the original story condition containing the twins' argument had a negative effect on subjects' gist recall.

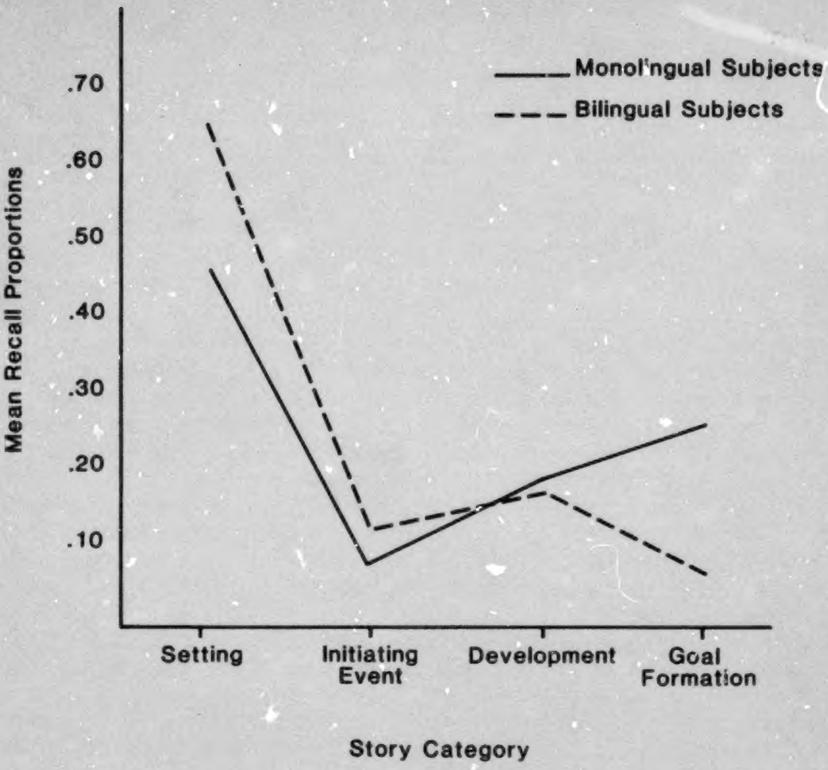


Figure 1. Mean recall by story category and sample in the original story.

The highly salient sibling argument in the setting category of the original story is consistent with van Dijk's (1979) notion of relevance assignment. He proposed that particular events in a selection may receive more processing attention, and therefore be more memorable because of various reader and text characteristics. In particular, the readers experience and the contrasting relationship of an event to its background seem to apply to the present case. That is, "The Chase Twins" is a fairly bland selection with a strong decoding emphasis. The characters are faceless and the dialogue mundane. However, children can relate to a family argument, especially when it is played against the dull backdrop of a quiet breakfast. Based on van Dijk's (1979, p. 125 "Tentative List of Relevance Signals in Discourse") and the results of this replication, future studies should consider the relative salience of particular propositions within the analysis of childrens' protocols. In order to verify the generalizability of propositional salience, future studies ought to employ multiple narratives. For example, two prose selections containing the twins' sibling argument but different surrounding context might further reveal the psychological validity of van Dijk's notion of relevance assignment or propositional salience.

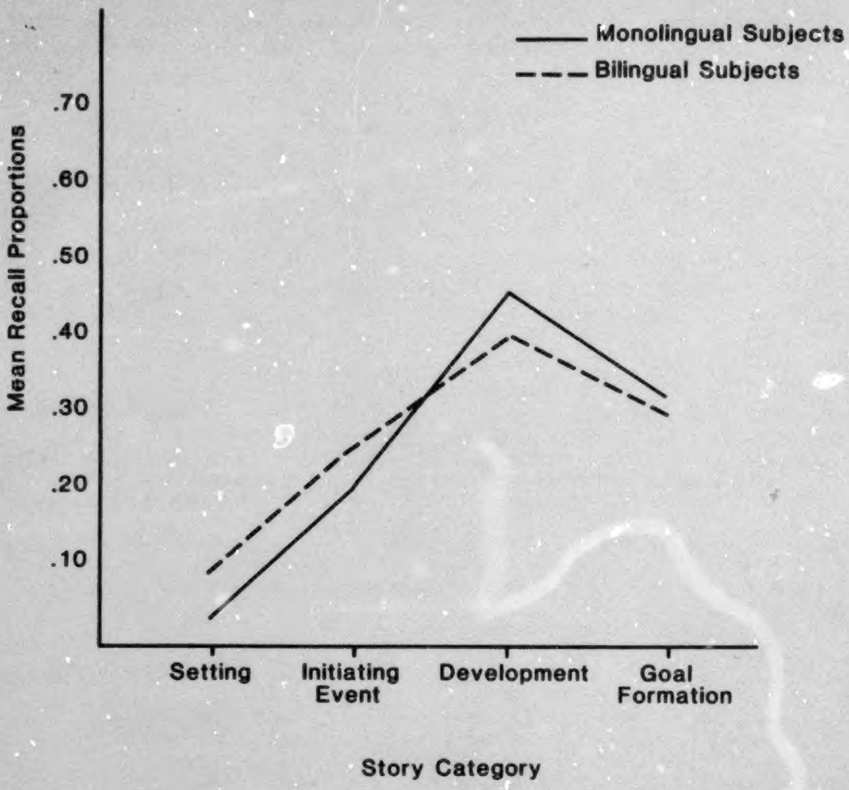


Figure 2. Mean recall by story category and sample in the story rewrite.

It may well be that propositional salience will become an important consideration in the development of instructional story material for children and in devising questions to stimulate understanding of story selections.

Finally, the divergent results achieved on the free versus probed recall measures may be considered in light of their varying task demands. The free recall measure may be an accurate reflection of the present subjects' comprehension but, due to language differences, less accurate for the bilingual group. Rigg and Taylor (1979) contrasted standardized measures of comprehension with retellings. They asserted that unlike the high risk, formal wh-question task, a free recall allows the reader as narrator to establish what is crucial in the story. However, this requires language fluency and confidence gained through experience with the retelling task. Both groups of students need opportunities to practice this activity with stimulating story material that is predictable and relevant to their lives.

In summary, the present study supports the previous (Bean et al., 1980) finding that middle grade students benefit from story material displaying semantic features of discourse that encourage comprehension as well as decoding ease.

These semantic features include a predictable story structure and propositions that are salient because they contribute to, rather than detract from childrens' comprehension.

ACKNOWLEDGEMENT

This research was funded in part by a faculty grant to the authors from California State University, Fullerton. The authors would like to thank Elaine McLaughlin for her assistance in scoring protocols. We would also like to acknowledge the helpful participation of Jim Soderberg, Vicki Flanagan, Jo Snyder, and Roger Walden, teachers at Vista Verde Elementary School and their students who made the study possible.

REFERENCES

Barnitz, J. G. Syntactic effects on the reading comprehension of pronoun-reference structures by children in grades two, four, and six. *Reading Research Quarterly*, 1980, *15*, 268-289.

Bean, T. W., Potter, T. C. & Clark, C. H. Selected semantic features of ESL materials and their effect on bilingual students' comprehension. In Michael Kamil & Alden Moe (Eds.), *Perspectives in Reading Research and Instruction.* Twenty-Ninth Yearbook of the National Reading Conference, 1980, 1-5.

Bower, G. H. Experiments on story comprehension and recall. *Discourse Processes*, 1978, *1*, 211-231.

Dreher, M. J. & Singer, H. Story grammar instruction unnecessary for intermediate grade students. *The Reading Teacher*, 1980, *3*, 261-268.

Mandler, J. M., A code in the node: the use of a story schema in retrieval, *Discourse Processes*, 1978, *1*, 14-35.

Mandler, J. M. & Johnson, N. S. Remembrance of things parsed: story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

Miami Linguistic Readers. Lexington, Massachusetts: D. C. Heath, 1966.

Nash-Weber, B. L. *Inference in an approach to discourse anaphora.* Urbana, Illinois: Center For The Study of Reading, Technical Report No. 77, 1978.

Nelson Reading Test. Boston, Massachusetts: Houghton Mifflin Co., 1977.

Raygor, A. L. The Raygor Readability Estimate: a quick and easy way to determine difficulty. In P. D. Pearson (Ed.), *Reading: Theory, Research, and Practice.* Twenty-sixth Yearbook of the National Reading Conference, 1977, 259-263.

Richek, M. A. Reading comprehension of anaphoric forms in varying linguistic contexts. *Reading Research Quarterly*, 1972, *2*, 145-165.

Rigg, P. & Taylor, E. Measures of comprehension: independence and implications. In Jerome C. Harste and Robert F. Carey (Eds.), *New Perspectives on Comprehension.* Bloomington, Indiana: School of Education, Indiana University, 1979, 77-83.

Rumelhart, D. D. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic Processes in Reading: Perception and Comprehension.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977, 265-303.

Stein, N. L. & Glenn, C. G. An analysis of story comprehension in elementary school children. In Roy C. Freedle (Ed.), *New Directions In Discourse Processing*. Norwood, New Jersey: Ablex Publishing Corporation, 1979, 53-120.

Stein, N. L. & Nezworski, T. The effects of organization and instructional set on story memory. *Discourse Processes*, 1978, *1*, 177-193.

Tannen, D. A comparative analysis of oral narrative strategies: Athenian Greek and American English. In Wallace L. Chafe (Ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production.* Norwood, New Jersey: Ablex Publishing Corporation, 1980, 51-87.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

Van Dijk, T. A. Relevance assignment in discourse comprehension. *Discourse Processes*, 1979, *2*, 113-126.

Warren, W. H., Nicholas, D. W., & Trabasso, T. Event chains and inferences in understanding narratives. In Roy C. Freedle (Ed.), *New Directions in Discourse Processing*. Norwood, New Jersey: Ablex Publishing Corporation, 1979, 23-52.

NANCY ROGERS-ZEGARRA
Wake Forest University

HARRY SINGER
University of Californa, Riverside

ANGLO AND CHICANO COMPREHENSION OF ETHNIC STORIES

Reading comprehension consists of an interaction between text-based and reader-based information (Rumelhart, 1977; Adams & Collins, 1977). An hypothesis drawn from this interaction view of reading comprehension is that when texts reflect only the Anglo experience, Chicanos are at a disadvantage and vice-versa. The purpose of this investigation was to test this hypothesis by assessing the comprehension of the two groups when (a) the text passages are drawn from both Anglo and Chicano experiences and (b) both groups are equated for general reading ability. More specifically, the hypothesis was tested that when Chicano students are equal in general reading ability to their Anglo peers, they have an advantage when assessed on inferences based on scriptal knowledge of Mexican-American culture and are at least equal when assessed on inferences based on scriptal knowledge of American culture. The reason the hypothesis may be tenable is that Chicanos participate in their own culture, of course, more so than Anglos do. However, Anglos, especially those in southern California, do become somewhat knowledgeable about Chicano culture. Therefore, they have some knowledge-base to draw upon for interacting with ethnic literature reflecting the Mexican-American culture. Chicanos also tend to experience the major cultural events in American culture as much as Anglos do and consequently have a knowledge-base for interacting with texts that is likely to equal Anglos' when reading stories based on significant events in the general American culture.

Previous comparisons of Anglo and Chicano students' reading abilities have been based upon standardized reading achievement tests. These tests attempt to minimize or eliminate assessment of text- and reader-based interactions. They especially try to avoid asking scriptal questions in testing comprehension; scriptal questions require the reader to draw upon background knowledge and use it in interaction with text-based information to arrive at correct answers (Pearson & Johnson, 1978). Indeed, some reader-based knowledge is necessary in reading printed materials of any kind because all the information necessary for comprehending a text passage is not contained within a text. Those readers who cannot mobilize the missing information from their knowledge repertoire are at a disadvantage (Marshall, 1976). This knowledge repertoire consists of three types of information: scripts, which are culturally determined sequences of events (Schank & Abelson, 1977); propositions, which are statements about the world (Kintsch, 1974; Winograd, 1972); and semantic maps, which consist of a concept with its properties, hyponyms, superordinates, and relationships with other concepts or semantic maps (Pearson & Johnson, 1978).

Because of the interaction between text-based and reader-based information, assessment of comprehension can vary on a continuum from text-based to reader-based interactions. Consequently, there can be a band of acceptable responses to comprehension of passages (Tierney & Spiro, 1979; Tierney & LaZansky, 1980). Although the interaction between text and prior knowledge is always present, requirements for use of prior knowledge may change according to (1) type of information being used, (2) purpose in reading, and (3) differences in individuals' processing levels (Spiro, 1979). What makes the interaction evident are responses to scriptal questions. Therefore, in this study the variations were type of questions (literal, implicit, and scriptal), type of information being read (Anglo and Chicano ethnic stories), and hence the degree to which prior knowledge must be drawn upon to interact with text-based information in answering the comprehension questions. Research related to these types of comparisons will be briefly reviewed.

REVIEW OF RELATED LITERATURE

Bartlett (1932) found that readers in England comprehended an American Indian folk-tale, "The War of the Ghosts," by modifying their recall and interpretation of events in the story so that the events *gradually* became more consistent with the schemata (knowledge structures) already present in the reader's background knowledge.

When readers come from different cultural backgrounds, they may interpret the same story differently *at the time of initial comprehension.* Steffenson, Joag-dev, and Andersen (1978) reported that East Indians and Americans interpreted an East Indian and an American wedding ceremony according to their own cultural backgrounds.

Even when readers come from the same cultural background, it is possible for them to interpret passages differently, depending on which high-level schemata are activated, as Pichert and Anderson (1977) and Anderson and Pichert (1978) demonstrated when they gave different perspectives to randomly selected college students, a "house buyer" perspective to one group and a "burglar" perspective to the other group.

Spontaneous selective storage of information from a text is likely to differ during learning sessions when individuals vary in their belief systems, as Levine and Murphy (1943) discovered. They had college students with pro- and anti-communistic social philosophies read the same passage dealing with the Soviet Union. The students read and reproduced the same passage twice over a time span varying from 15 minutes to 4 weeks. The rate of learning the anti-Soviet passage was greater for the anti-communistic students, but the decline in rate of retention for the different amounts of information each group had learned was about the same for both groups when tested weekly over a five-week period.

Since cultures are not homogeneous, even within a nation, groups of individuals can be identified whose perspectives on issues and events will vary considerably. Consequently, their scriptal or reader-based comprehension of the same passage will differ. One such contrast is urban vs. rural children. Reynolds, Taylor, Steffenson, Shirey, and Anderson (1980) found that at the eighth-grade level differences in comprehension of urban students, white and black, when compared with rural white students were related to variation in the students' background knowledge for a story. Goodman (1978) also concluded that variation in prior knowledge and experience accounted for differences in reading comprehension of groups of elementary school children who were non-native and dialectal speakers drawn from four different cultural backgrounds.

Cultural differences can occur not only for interpretation of stories, but also for particular words. These variations in interpretation can have dramatic effects in everyday life. For example, Chicanos interpreted medical benefits for the "family" as including not only the nuclear but also the extended family. Consequently, their extended family showed up for medical treatment, which led their employers to a hasty modification of the wording of the contract for medical benefits. Thus, some of the differences in reading comprehension between Chicanos and their Anglo peers, even for text-based literal and inferential comprehension, can be attributed to differences in cultural perspectives and background knowledge.

But none of the cross-cultural studies reviewed here equated the comparison groups on general reading ability. Moreover, none of the studies attempted to determine whether the two comparison groups were equal in text-based literal and inferential comprehension. These methodological limitations were rectified in the design of this study.

EXPERIMENTAL DESIGN

Independent Measures

A total of 173 students, consisting of two ethnic groups, Anglos and Chicanos, drawn from grades 5 and 8 from four schools located in Riverside, California, participated in the study. Two grade levels were used because students' background knowledge of American and Mexican-American cultures should improve with age. Moreover, the older students, Anglo and Chicano, should have a greater degree of knowledge about *both* cultures.

At each grade level, each of the two ethnic groups was equated on standardized tests of reading achievement. Since the standardized reading tests emphasize text-based literal and implicit types of comprehension, the two groups should be equal when assessed on literal and implicit types of comprehension for ethnic stories. Each group was then further subdivided into two groups of readers: high (two years on the average above grade level) and low (two years on the average below grade level).

Since previous research has shown that students from higher socio-economic backgrounds perform better on comprehension tests, the readers were also classified according to two socio-economic levels on the occupational index, high (professionals) versus low (non-professionals).

Thus, the study employed a 2 × 2 × 2 × 2 experimental design. It consisted of two ethnic groups (Anglos versus Chicanos), two grade levels (5 and 8), two reading levels (high versus low), and two socioeconomic levels (high versus low).[1]

Dependent Measures

Thirty-six multiple-choice questions were constructed for three levels of comprehension, following Pearson and Johnson's criteria (1978). Answers to the literal questions are provided in the story. Answers to the implicit questions can be inferred from text-based information. However, the scriptal questions require the reader to use prior knowledge in interaction with text-based information in order to arrive at correct answers. Thus, there were three dependent measures, consisting of twelve comprehension questions for each type of comprehension. These questions were equally divided among the two types of ethnic stories.

---

[1] The data used in this report are part of the information collected for a more detailed investigation (Rogers-Zegarra, 1981)

Materials: Ethnic Stories

Twelve stories were written, six based on Mexican culture and six on Anglo culture. The Mexican culture stories concerned holidays and customs unique to Mexican-Americans in the Southwestern United States (e.g., Las Posadas or El dia de los Muertos). The Anglo stories dealt mainly with holidays or customs unique to North American Anglo life (e.g., the Fourth of July celebration and April Fool's Day).

All of the stories were written at approximately an equal level of difficulty. The range of readability varied from the fourth-grade to the seventh-grade level of difficulty as assessed by the Fry readability formula. This range of variation is within the formula's standard error of measurement. The stories ranged in length from 250 to 1,000 words, but were balanced for length for both groups.

Four adult representatives, two Anglos and two Chicanos, rated the stories for ethnic content. Each pair of raters agreed that the stories consisted of content that was familiar.

Procedure. All the stories were read aloud to the students. Each student also had a copy of each story and therefore could follow along, reading silently. Thus, any word recognition difficulties were minimized, while comprehension was maximized (Goldstein, 1940).

RESULTS

For *literal questions*, the analysis of variance was significant only for reading level ($F = 31.97$, df 1,157, $p < .001$). The better readers at both grade levels were superior at recalling or locating information directly stated in the text.

For *implicit questions*, significant differences were found for grade ($F = 25.07$, df 1,157, $p < .001$) and reading level ($F = 25.29$, df 1,157, $p < .001$). These results were also expected because older students are generally better in recalling information and at making inferences from text-based information.

For *scriptal questions*, the results were again significant for grade ($F = 19.16$, df 1,157, $p < .001$) and reading levels ($F = 12.73$, df 1,157, $p < .001$). The older and better readers were not only superior in drawing inferences, but also had greater background knowledge to draw upon for answering scriptal questions. However, a trend for ethnicity ($F = 2.70$, df 1,157, $p = .10$) and a trend toward interaction ($F = 3.72$, df 1,157, $p < .056$) for ethnicity and reader level occurred. The better readers among the Chicano students, regardless of grade level, had significantly higher scriptal comprehension for text involving content based on Chicano culture than did the Anglo high and low readers and the Chicano low readers. Chicano students also performed as well as Anglos on scriptal questions for Anglo literature. These results indicate that the Chicanos have bicultural knowledge to draw upon in reading either type of literature. This knowledge is superior to Anglos for answering scriptural questions on the Chicano literature and equal to Anglos for answering questions on the Anglo literature.

Socio-economic level was not statistically significant in any of the analyses, probably because equating the groups on reading level and using professional level as a dividing point for high versus low socio-economic levels washed out differences attributable to socioeconomic levels.

SUMMARY AND IMPLICATIONS

The results of this study indicated that when Chicanos and Anglos from grades 5 and 8 are matched on general reading ability as assessed by standardized tests the older and better readers from both ethnic groups were signficantly better than the younger readers on all types of comprehension. Within each grade level, the Chicanos were equal to the Anglos on text-based literal and implicit types of comprehension, but the Chicanos, and particularly the better readers among the Chicanos across grade levels, were superior to the Anglos on scriptal questions for Anglo-based stories.

This study indicates that in order to have a comprehensive assessment of reading ability, it is necessary to measure all three types of comprehension: literal, implicit, and scriptal. The study also demonstrates the importance of assessing students on different types of ethnic literature because when Chicanos are equated with Anglos on general reading ability, those students with a bicultural background tend to perform as well on text-based comprehension and can exhibit their advantage in text- and reader-based interactions over their Anglo peers for Chicano-based literature. At the same time, Anglos who have the opportunity to read multiethnic literature and participate in the non-Anglo culture also tend to expand their background knowledge and are better able to answer scriptal questions, as shown by the better performance of the older over the younger students for both types of literature. This expansion occurs in Southern California where events for both Anglo and Chicano cultures are celebrated in the schools and in the community. Thus, both groups benefit from development of bicultural schemata in many ways and particularly in comprehension of Anglo and Chicano ethnic stories.

REFERENCES

Adams, M. J. & Collins, A. *A schema-theoretic view of reading* (Technical Report No. 32). Urbana, Illinois: Center for the Study of Reading, University of Illinois, 1977.

Anderson, R. C., & Pichert, J. W. Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior,* 1978, *17,* 1-12.

Bartlett, F. C. *Remembering.* Cambridge: The University Press, 1932.

Goldstein, H. *Reading and listening comprehension at various controlled rates* (Teachers College Contributions to Education, No. 821). New York: Bureau of Publications, Columbia University, 1940.

Goodman, K. S. *Reading of American children whose language is a stable rural dialect of English or a language other than English* (Final Report). Washington, D.C.: National Institute of Education, U.S. Department of Health, Education and Welfare, 1978.

Joag-dev, C., & Steffensen, M. S. *Studies of the bicultural reader: Implications for teachers and librarians* (Reading Education Report No. 12). Urbana, Illinois: Center for the Study of Reading, University of Illinois, 1980.

Kintsch, W. *The representation of meaning in memory.* Hillsdale, New Jersey: Erlbaum-Wiley, 1974.

Levine, J. M., & Murphy, G. The learning and forgetting of controversial material. *Journal of Abnormal and Social Psychology,* 1943, *38,* 507-517.

Marshall, N. *The structure of semantic memory for text.* Unpublished doctoral dissertation, Cornell University, 1976.

Pearson, P. D., & Johnson, D. D. *Teaching reading comprehension.* New York: Holt, Rinehart and Winston, 1978.

Pichert, J. W., & Anderson, R. C. Taking different perspectives on a story. *Journal of Educational Psychology,* 1977, *69,* 309-315.

Reynolds, R., Taylor, M., Steffensen, M. S., Shirey, L., & Anderson, R. C. *A study of the effect of* *cultural knowledge on reading comprehension: A comparison of black and white*. Urbana, Illinois: The Center for the Study of Reading, University of Illinois, 1980.

Rogers-Zegarra, N. *Anglo and Chicano literal, implicit and scriptal comprehension of ethnic stories*. Doctoral dissertation in progress, University of California, 1981.

Rumelhart, D E. Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, New Jersey: Erlbaum, 1977.

Spiro, R. *Prior knowledge and story processing: integration, selection, and variation* (Technical Report No. 138). Urbana, Illinois: Center for the Study of Reading, University of Illinois, 1979.

Steffensen, M. S., Joag-dev, C., & Anderson, R. C. *A cross-cultural perspective on reading comprehension* (Technical Report No. 97). Urbana, Illinois: Center for the Study of Reading, University of Illinois, 1978.

Tierney, R. J., & LaZansky, J. The rights and responsibilities of readers and writers: A contractual agreement. *Language Arts*, 1980, *57*, 606-613.

Tierney, R. J., & Spiro, R. J. Some basic notions about reading comprehension: Implications for teachers. In J. Harste and R. F. Carey (Eds.), *New perspectives on comprehension* (Monograph No. 3 in Language and Arts Series). Bloomington, Indiana, 1979.

Winograd, T. Understanding natural language. *Cognitive Psychology*, 1972, 3 (whole issue).

ALLAN R. NEILSEN
University of Calgary

AN INVESTIGATION OF THE RELATIONSHIP OF COHESION TO LINGUISTIC MARKING, DISCOURSE STRUCTURE AND CONTENT FAMILIARITY

PURPOSE AND BACKGROUND

The purpose of this exploratory study was to investigate the relationship of cohesion (Halliday and Hasan, 1976) to linguistic marking, discourse structure, and content familiarity in order to help clarify their effect on the comprehension and recall of written discourse.

A study conducted by Neilsen (1977) within this general framework examined the effect of linguistic marking of logical relationships in text (explicit versus implicit), discourse structure (hierarchical versus sequential), and content familiarity (familiar versus unfamiliar) on reading comprehension and recall of text. Results indicated that, although differences in discourse structure and content familiarity significantly influenced both comprehension and recall, linguistic marking did not. The finding for linguistic marking was somewhat surprising since it was anticipated that the explicit marking of logical relationships would aid comprehension and recall. It was speculated that the finding might have been due to the fact that only good readers were involved and that they were able to make the inferences necessary to reconstruct meaning. In fact, there is some evidence (Marshall and Glock, 1978-79) to indicate that this might be the case. Further, it was suggested that general semantic constraints within the passage rendered explicit marking redundant. The purpose of this study was to clarify the effect of linguistic marking. The original investigation was expanded to include examination of the relationship between semantic constraints and discourse structure and content familiarity. Halliday and Hasan's (1976) notion of cohesion provided the basis for analyzing the semantic properties of the original experimental passages. The following question guided the investigation:

> Does cohesion covary reliably with differences in linguistic marking, discourse structure or content familiarity in a manner that might affect the comprehension or recall of written text?

COHESION

Cohesion is a semantic concept that refers to the relations of meaning that exist within a text. Cohesion occurs where the *interpretation* of one element in text is dependent on its relationship to another, usually presupposed or preceding, element. Halliday and Hasan use the term *tie* to refer to a single instance of cohesion (i.e. occurrence of a pair of cohesively related items). For example, in (1) the relationship between "farmers" and "they" is cohesive and constitutes one tie:

(1) There are many *farmers* in the country.
They are very poor.

In this example, "they" is referred to as the *cohesive item* and "farmers" as the *presupposed item*. Thus a tie is comprised of one cohesive item and one presupposed item. It is important to note that, since cohesive ties are intersentential in nature, the cohesive and presupposed items cannot occur within the same sentence.

The concept of tie allows texts to be characterized in terms of the number of cohesive relations they contain and thus provides one index of cohesiveness.

Halliday and Hasan indicate that although cohesion is a semantic relation it is realized through the lexico-grammatical system. Of the five main types of cohesive ties delineated, four (reference, substitution, ellipsis, conjunction) are realized through the grammar, while one (lexical) is realized through the vocabulary. Only reference, conjunction, and lexical ties occurred with sufficient frequency in the passages examined in this investigation to warrant further description and analysis. A brief explanation and example of each follows:

(2) *Reference:* cohesive items that cannot be interpreted in their own right. They are usually pronouns and thus are entirely dependent on their relation to presupposed items for interpretation.
e.g. The *animals* were left homeless. *They* were angry.

(3) *Conjunction:* cohesive items that signal very general underlying semantic relations (e.g. causal, temporal, adversative, etc.)
e.g. *Circle Island has good soil. However,* there are few rivers and therefore there is a shortage of water.

(4) *Lexical:* cohesive items that are identical or hierarchically related to a preceding lexical item.
e.g. The farmer pulled the *donkey*. The *donkey* wouldn't move.

This delineation of types of cohesive relationships makes possible a qualitative analysis of texts. For instance, the relative distribution of different types of ties might provide another means for describing differences in cohesiveness among texts.

In addition to number and types of cohesive relations in texts, Halliday and Hasan's system also provides for the analysis of cohesive distance. Cohesive distance is determined by the number of sentences intervening between the cohesive item and the presupposed item in the text. For example, in (5) the cohesive distance is 0 because the cohesive elements appear in contiguous sentences. On the other hand, the cohesive distance between the elements in (6) is 2 because there are two sentences between the cohesive item ("union") and the presupposed item ("union").

(5) The farmer pulled the *donkey*. The *donkey* wouldn't move. (Cohesive distance = 0).

(6) The Whigs asked the weathermen's *union* to support the bill. The Tories would join too. The bill would win passage in the senate. The *union* would not support the bill. (Cohesive distance = 2).

Cohesive distance adds a sense of the extent to which cohesive elements are dispersed within a text and hence provides still another dimension for the characterization and comparison of text cohesiveness.

INDICES OF COHESION

Based on the concepts of ties and cohesive distance, three main indices of cohesion were derived.

Density of Ties

Density of ties, measured by the average number of ties per sentence, provides an index of the relative amount or degree of cohesion in a text. Presumably, the greater the number of ties, the more cohesive the text. Further, it seems reasonable to suggest that as the number of ties in a text increases, the easier it will be to construct and remember because the semantic relationships are more highly integrated than those in a text which contains fewer ties (i.e., is less cohesive).

Cohesive Distance Per Tie

Average cohesive distance per tie provides an index of the general dispersion of ties within a text. As the distance between cohesive elements increases, linking the elements to form a tie may require more effort. A text with relatively long distance between cohesive elements might be considered less cohesive and thus more difficult to understand and remember than a text in which average cohesive distance is relatively short.

Distribution of Types of Ties

Some insight regarding qualitative differences in texture or cohesiveness can be gained by examining the relative proportion of different cohesive ties across texts. Intuitively, it seems unlikely that all types of ties will be of equal "strength". Some will be more difficult to process. This might affect unity or overall cohesiveness which, in turn, might affect comprehensibility and memorability.

Reference. Due to their relative abstractness and total dependence on presupposed elements for their interpretation, reference ties might place great demands on attention. In fact, a high proportion of reference ties, although contributing to cohesion technically, might actually result in ambiguity of relationships within a text.

Conjunction. Since conjunctive ties underscore or make explicit logical relations underlying text, a high proportion of such ties should contribute to the overall "tightness" of text and facilitate understanding and recall.

Lexical. Lexical ties involve matching elements (words) that are identical, synonymous, or generically related. Therefore, a relatively high proportion of lexical ties should facilitate text processing and recall.

ANALYSIS

The eight passages analyzed in this investigation embodied unique combinations of linguistic marking, discourse structure, and content familiarity. In order to obtain the cohesion indices, the number, type, and cohesive distance was calculated for each sentence in all passages. Next, the density of ties was derived by dividing the total number of ties per passage by the number of sentences in the passage. The average cohesive distance per tie was derived by dividing the total distance by the number of ties in the passage. The distribution of different types of ties was calculated by dividing the total number of instances of a particular type of tie in a passage by the total number of ties in the passage.

Next, results for each block of passages representing a single variable (e.g. Linguistic Marking) were assembled. Results across levels within each variable (e.g. Explicit vs Implicit) were then compared. The Mann-Whitney U text (Siegel,

1956) was used to compare results for the average number of ties per sentence, the average distance per tie, and the distribution of different types of ties.

RESULTS AND INTERPRETATION

The purpose of this investigation was to determine whether cohesion covaries reliably with differences in linguistic marking, discourse structure, or content familiarity in a manner that might affect comprehension and recall of written text. The relationship of cohesion to each of these variables is discussed in terms of the three indices described above. Results of the cohesion analysis are also used to help clarify the findings from the original study (Neilsen, 1977).

Cohesion and Linguistic Marking

The Mann-Whitney U test revealed that the density of ties in explicit passages as significantly higher ($p = <.014$) than in implicit passages. This finding was not surprising since implicit passages systematically omitted all conjunctive ties. This, combined with the finding that cohesive distance was less ($p = <.057$) for explicit passages, suggests that the explicit passages were more cohesive. On the basis of these findings, it would be reasonable to predict that comprehension and recall for explicit passages would be better than for implicit passages. However, actual findings from the original experiment showed no difference between the two marking conditions. Thus, although cohesive density and distance did vary quantitatively with linguistic marking, they did not affect reading performance.

The difference in density of ties can be attributed directly to the greater number of conjunctive ties in explicit passages. Halliday and Hasan (1976) explain that although conjunction might make semantic relations more explicit, it is not a necessary condition for apprehending and comprehending these relations.

This explanation might account for the lack of difference in comprehension and recall, particularly since only proficient readers were involved. Thus, it seems important to consider both qualitative and quantitative dimensions when assessing the influence of cohesion on text processing. A further qualitative consideration is the high proportion of lexical ties (due mainly to extensive repetition of words and phrases) in both explicit and implicit passages. It is possible that the cohesive power of these ties was sufficient to neutralize differences in marking (i.e., conjunctive ties).

The difference in average cohesive distance can be seen as an artifact of omitting conjunctive ties in the implicit passages which resulted in the proliferation of simple sentences and hence increased cohesive distance. In fact, the effect of distance may have been confounded by the demands of processing additional chunks of information.

Cohesion and Discourse Structure

Cohesion did not vary quantitatively across discourse structure. Both hierarchical and sequential structure were similar in terms of density and average cohesive distance of ties. This suggests that comprehension and recall for both types of passages might also be similar. Again, actual findings fail to support this prediction. Performance on hierarchically structured passages was consistently and significantly better than on sequentially structured passages. Since cohesion does not appear to have been a factor, differences in performance can be attributed to differences in structure. This interpretation must be qualified by reference to the difference in proportion of reference ties between structure types. Sequential passages had a significantly greater ($p = < .029$) percentage of reference ties than hierarchical passages. Since reference ties are believed to be more difficult to process, they might have contributed to the performance differences attributed to structure alone.

Cohesion and Content Familiarity

Results of the cohesion analysis revealed no significant difference in the density of ties in familiar and unfamiliar content. Despite this similarity in density of ties, comprehension and recall of familiar content was consistently and significantly better than for unfamiliar content. This indicates that density of ties did not affect performance. However, it appears as though the average distance per tie might have contributed to differences in performance. If it can be reasoned that an increase in the average distance between cohesive elements would make the integration and recall of meaning more difficult, then the greater ($p = < .057$) average distance in unfamiliar content passages might have contributed to the performance differences attributed solely to differences in abstractness of lexical content. In fact, this finding suggests that cohesive distance might be a dimension of "familiarity". Examination of results for proportion of different types of ties does not indicate any variation in cohesion across familiarity levels and therefore does not allow any further qualification on the original interpretation of findings for content familiarity.

CONCLUSIONS

Based on the results of this investigation, the following "conclusions" are tentatively drawn:

1. density of ties and average cohesive distance seem to vary reliably with differences in linguistic marking of text but do not seem to affect comprehension or recall.
2. neither cohesive density nor average cohesive distance seems to vary reliably with differences in discourse structure and therefore do not seem to be factors that affect comprehension and recall. However, a high proportion of reference ties in sequentially structured passages might have contributed to the significantly lower performance scores for these passages in the original study.
3. while density of ties does not appear to be a dimension of content familiarity that affects comprehension or recall, average cohesive distance might be. Greater average difference between ties in unfamiliar content may have contributed to lower scores on these passages.

Due to the speculative nature of the findings and conclusions, there seems to be no reason to alter the original interpretation regarding the effect of linguistic marking, discourse structure, or content familiarity. However, the concept of cohesion does provide another dimension to consider when analyzing the demands that texts place on readers. Additional investigation with a much larger number of texts varying on the dimensions included in this study is necessary in order for any trends in the effect of cohesion variation to emerge and stabilize.

REFERENCES

Halliday, M. and Hasan, R. *Cohesion in English*, Longman, 1976.

Marshall, N. and Glock, M. Comprehension of connected discourse: A study into the relationship between the structure of text and information recalled. *Reading Research Quarterly, 14*, 1978-79, 10-56.

Neilsen, A. *The role of macrostructures and linguistic connectives in comprehending familiar and unfamiliar written discourse*. Unpublished doctoral dissertation, University of Minnesota, 1977.

Siegel, S. *Nonparametric statistics for the behavioral sciences*. McGraw-Hill, 1956.

GREGORY G. MCKENZIE
Calgary Board of Education

ALLAN R. NEILSEN
CARL BRAUN
University of Calgary

THE EFFECTS OF LINGUISTIC CONNECTIVES AND PRIOR KNOWLEDGE ON COMPREHENSION OF GOOD AND POOR READERS

The purpose of the study was to examine the effects of linguistic connectives and prior knowledge on the comprehension of causal relationships in written discourse by good and poor readers.

Rumelhart's (1975) widely accepted model of the reading process suggests that a skilled reader employs both conceptually-based and textually-based sources of information concurrently when attempting to make sense of the printed page. Evidence from a number of studies investigating the conceptual domain suggests that prior knowledge/experience is a critical factor in the interpretation of text (Anderson, Reynolds, Schallert and Goetz, 1976; Bransford and Johnson, 1972; Pichert and Anderson, 1976). Related research has confirmed the intuitively sensible notion that prior knowledge of topic facilitates comprehension and recall of what is read (Mandler and Johnson, 1977; Marshall and Glock, 1978-79; Neilsen, 1977; Pace, 1978; Pearson, Hansen and Gordon, 1979.)

In the realm of text-based information, there is a body of literature on the effect of linguistic connectives on comprehension. However, little of the research has involved written text and even less has addressed the effect of connectives in extended written discourse. Further, it seems that some of the more common relationships signalled by linguistic connectives have received little attention. There are some studies that suggest the presence of connectives facilitates comprehension (Marshall and Glock, 1978-79; Pearson, 1974-75). However, several others (Hawley, 1979; Neilsen, 1977) found no difference in comprehension due to the presence or absence of connectives.

The present study seemed warranted for several reasons. First, in most previous studies, neither the study of linguistic connectives as cues to logical relationships in written text, nor the study of prior knowledge were submitted to direct investigation.

Second, despite the high frequency of their appearance in print (Carroll, Davies and Richman, 1971; Robertson, 1968; Rogers, 1975), few studies have investigated the effect of connectives used to signal causal relationships.

Third, most previous investigations of the role of connectives have been conducted at the word and sentence levels (Bart, 1974; Gardner, 1974; Pearson, 1974-75; Robertson, 1968; Stoodt, 1970) rather than at the level of connected discourse.

Finally, few studies (Marshall and Glock, 1978-79; Zinn, 1979) have probed the possible interactions of linguistic marking and prior knowledge with reading ability.

Three specific questions guided the investigation:

1. Do linguistic connectives affect the comprehension of causal relationships in written discourse?

2. Does content familiarity affect the comprehension of causal relationships in written discourse?

3. Do linguistic connectives, content familiarity, and reading ability operate inter-dependently in the comprehension of causal relationships in written discourse?

DESIGN AND ANALYSIS

The experiment employed a 2 × 2 × 2 factorial design with reading ability (Good and Poor Readers), and linguistic connectives (Explicit and Implicit), as between-subject variables and content familiarity (Familiar and Unfamiliar) as the within-subject variable.

Good Readers were operationally defined as students whose Paragraph Comprehension scores on the Nelson Reading Tests (1962) were in stanines six or seven while Poor Readers were defined as those with scores in stanines three or four. Four experimental passages were constructed. In the Explicit versions, the connective "because" was included to signal each of four target causal relationships within each passage. In the Implicit versions, "because" was omitted. Two Familiar and two Unfamiliar passages were selected on the basis of extensive pilot work with a sample of students representing the same characteristics as those in the main study. On the basis of results derived from four different indices (wh-probes, word association, passage ranking, cloze). The topics *Obtaining a Drivers License* and *Energy Sources* were identified as Familiar. Unfamiliar topics were *Energy Past* and *Navigation*. Readability was "controlled" separately from content familiarity in an attempt to minimize comprehension difficulties caused primarily by unknown lexical items and complex syntax. The Fry (1968) formula was used to write passages for Good Readers at the grade seven level and passages for Poor Readers at the grade four level. This compromise was made to avoid the ceiling effects often obtained when using low readability passages with good readers.

Comprehension of the target causal relationships as measured by wh-probes was the dependent variable.

A three-way analysis of variance with one repeated measure was used to analyze the main effects. Scheffé post hoc comparisons were used to clarify the nature of the interactions.

The experimental sample (N = 68, 17 per cell) was drawn from the population of grade seven students in two middle class area schools.

PROCEDURES

Half the students read the Explicit passages. The other half read the Implicit versions. Familiar passages were randomly interspersed with the Unfamiliar passages within each group.

Students were asked four questions each for the two familiar and two unfamiliar passages read and were allowed to look at the passages when writing their answers. Answers were given a value of 1 if correct or 0 if incorrect. Hence, a perfect score across all four passages was 16. All answers were scored by the main investigator and 20 percent of the papers were also scored by an independent marker with 99 percent agreement.

RESULTS

The only significant main effect ($F=25.32$; $p=<.001$) was for content familiarity favoring Familiar ($\bar{X}=4.79$, S.D. = 1.61) over Unfamiliar ($\bar{X}=3.71$, S.D. = 1.82). The following post hoc findings indicated that there may be some interdependence between reading ability, linguistic connectives, and content familiarity:

1. Good Readers ($\bar{X}=5.65$, S.D. = 1.00) performed significantly better ($F=26.82$; $p<.01$) than Poor Readers ($\bar{X}=3.41$, S.D. = 1.28) on Implicit versions of Familiar passages.
2. Poor Readers performed significantly better ($F=15.63$; $p<.05$) on Explicit versions of Familiar passages ($\bar{X}=5.59$, S.D. = 1.99) than on Explicit versions of Unfamiliar passages ($\bar{X}=2.88$, S.D. = 1.87).

DISCUSSION

On the basis of post hoc results, the study tentatively suggests that the differences in comprehension between Good and Poor Readers may not be uniform across Explicit/Implicit versions and Familiar/Unfamiliar content. The finding that Good Readers appeared to comprehend significantly better than Poor Readers in implicitly marked versions of Familiar content may indicate that implied relationships are more difficult for Poor Readers to understand due to the additional cognitive and, perhaps, metacognitive processing required in making inferences. Further, it seemed that Poor Readers comprehended Explicit passages that were Familiar significantly better than those that were Unfamiliar. It may be that the Unfamiliar content proved so overwhelming that even when passages were presented in Explicit form, comprehension was not significantly enhanced. Poor Readers were not able to capitalize on the textual cues (linguistic connectives) provided. This finding is consistent with the main effect reported for content familiarity in this study and underscores the potency of prior knowledge in comprehending written discourse as discussed by Anderson et al (1976), Neilsen (1977) and Pearson, Hansen and Gordon (1977).

IMPLICATIONS

From a theoretical point of view, post hoc findings imply that processing in reading may be more difficult for poor readers than for good readers. The possible interdependence of reading ability, linguistic connectives, and content familiarity suggest this. Good readers may cope relatively well with both conceptual and textual processing. Poor readers however, may sometimes cope with either textual or conceptual difficulties but not always with both. In the present study, when unfamiliar and/or implicit conditions were present, the magnitude of loss in comprehension was significantly greater for poor readers than for good readers. More specifically, these findings imply that the ability to make inferences about logical relationships represented in text is one dimension that differentiates good from poor readers. This in turn suggests that explicit teaching of selected writing conventions (e.g., use of connectives) and the development of inferencing strategies might be helpful to poor readers.

The study also contributes to the growing body of evidence that indicates the importance of considering students' probable conceptual development and experiential background when developing/selecting instructional materials.

Finally, given that prior knowledge appears to be an exceptionally potent factor, it seems imperative that its influence be controlled or manipulated systematically in any future experimental investigation of factors involved in comprehending written discourse.

REFERENCES

Anderson, R. C., Reynolds, R. E. Schallert, D. L. & Goetz, E. T. *Frameworks for comprehending discourse* (Tech. Rep. 12), Urbana, Il.: University of Illinois Laboratory for Cognitive Studies in Education, July 1976.

Bart, W. Adolescents' understanding of logical connectives. *Psychological Reports*, 1974, *35*, 41-42.

Bransford, J. & Johnson, M. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behaviour*, 1972, *11*, 717-726.

Carroll, J. B., Davies, P. & Richman, B. *Word frequency book:* Boston: Houghton Mifflin, 1971.

Fry, E. B. A readability formula that saves time. *Journal of Reading*, 1968, *11*, 513-515; 575-578.

Gardner, P. L. *Logical connectives in science. A preliminary report:* Paper presented at the Sixth annual Conference of the Australian Science Education Research Association, May 1975.

Hawley, P. *The effect of prior knowledge and linguistic connectives on comprehension of causal relations in written discourse:* Unpublished master's thesis, University of Calgary, 1979.

Mandler, J. & Johnson, N. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

Marshall, N. & Glock, M. Comprehension of connected discourse: A study into the relationship between the structure of text and information recalled. *Reading Research Quarterly*, 1978-79, *14*, 10-56.

Neilsen, A. R. *The role of macrostructures and linguistic connectives in comprehending familiar and unfamiliar written discourse:* Unpublished doctoral dissertation, University of Minnesota, 1977.

Nelson, M. J. *The Nelson Reading Test.* New York: Houghton Mifflin, 1962.

Pace, A. *The influence of world knowledge and metacomprehension ability on children's comprehension of short narrative passages:* Unpublished doctoral dissertation, University of Delaware, 1978.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall and conception of certain semantic relations. *Reading Research Quarterly*, 1974-75, *10*, 155-192.

Pearson, P. D. Hansen, J. & Gordon, C. *The effect of background knowledge on young children's comprehension of explicit and implicit information*, (Tech. Rep. 116), Urbana, Il.: Center for the Study of Reading, March 1979.

Pichert, J. & Anderson, R. C. *Taking different perspectives on a story* (Tech. Rep. 91), Urbana, Il.: Center for the Study of Reading, November 1976.

Robertson, J. Pupil understanding of connectives. *Reading Research Quarterly*, 1968, *3*, 387-417.

Rodgers, D. Which connectives? Signals to enhance comprehension. *Journal of Reading*, 1974, 17, 462-466.

Rumelhart, D. *Toward an interactive model of reading*, Unpublished manuscript, University of California, San Diego, 1976.

Stoodt, B. *The relationship between understanding grammatical conjunctions and reading comprehension*, Columbus, Ohio: Ohio State University, September 1970.

Zinn, C. *A comparison of high and low grade four readers on their use of the logical connective "because":* Unpublished master's thesis, University of Alberta, 1979.

TESS O. KOSOFF
Syracuse University

THE EFFECTS OF THREE TYPES OF WRITTEN INTRODUCTIONS ON CHILDREN'S PROCESSING OF TEXT

A number of recent research studies have indicated that events which occur immediately prior to reading can be "powerful determinants" of the comprehension and retention of a text (E. Gagne, 1975, p. 645). Investigations have been conducted on the effect of such prior events as thematic titles (Bransford and McCarrell, 1974; Dooling and Lachman, 1971; Dooling and Mullet, 1973; Schallert, 1976), thematic drawings (Bransford and Johnson, 1972), and advance organizers (Ausubel, 1960; Mayer, 1979). Within schema theory, it is postulated that such prior events operate to activate the reader's relevant knowledge structures, such that the set of activated knowledge structures, or schemata, serve as a conceptual framework, or context, as the reader interacts with a subsequent text (Schallert, 1976). Moreover, there is evidence that the prior presentation of different information—different titles (Schallert, 1976), different background settings (Brown, Smiley, Day, Townsend, and Lawton, 1977), or different assigned perspectives (Pichert and Anderson, 1977)—induces differences in subjects' recalls of a given stimulus text. Thus, these studies strongly suggest that different types of introductions are likely to activate qualitatively different schemata for processing an immediately following passage and consequently leading to different reader-text interactions.

Since most of the research studies addressing the effects of textual introductions have utilized advance organizers and some sort of control passage, there has been little attempt to study the effect of other kinds of introductions: introductory material representative of the sort teachers might devise for preparing students to read new basal selections, introductions such as those suggested in teachers' manuals for basal series, or the types of written introductions that sometimes precede specific selections in current basal readers (e.g., Holt Basic Readers, 1977; Keys to Reading, 1975; Reading Basics Plus, 1976). The present study sought to investigate the effects of three qualitatively different types of written introductions, presented to children immediately prior to their reading an informational passage, on the composition of the knowledge structures constructed in the interaction with the stimulus passage. It was expected that there would be differences in the composition of the children's knowledge structures, as reflected in their immediate and delayed written recalls, as a function of the different types of introductions or the absence of an introduction. Written free recall was used as the basis for assessing students' comprehension and learning from text because this type of recall is free from the contaminating effects of externally provided cues (Shimmerlik, 1978) and also because Brown and Smiley (1977) demonstrated that no significant differences existed in the amounts of information (measured in idea units) contained in the written and oral free recalls of fifth grade students.

METHOD

Subjects

Subjects were fifth grade students from two schools in the same suburban school district near Syracuse, New York. All fifth graders in the two schools, who had achieved grade equivalent scores of 5.0 or above on a standardized reading achievement test (administered in Fall, 1979), participated in the study. After data collection, a total of 16 students were randomly eliminated to obtain a sample of 112 students such that students were equally distributed within school, treatment condition, and passage.

Materials

The materials for each of the two testing sessions were compiled into individual booklets for students to complete in a self-paced manner.

Stimulus passages. Seven, non-narrative passages on social science topics, from *Reading for Concepts* (Book C) (Liddle, 1970), were slightly modified to meet criteria set for length (165–175 words), total number of propositions (30–35), and a fourth grade readability level, according to the Fry formula. To control for content familiarity, it was priorly ascertained that none of the passage topics was specified in the students' science and social science curricula. To further control for the effects of students' prior knowledge about the passage topics, students were randomly assigned to the stimulus passages.

Introductions. Three types of introductions were constructed to accompany each of the seven passages. Following the Pearson and Johnson (1978) taxonomy of question-answer relations, the three types of introductions were constructed on the basis of the relationship that the introduction held to the passage. A textually explicit introduction contained information that was explicitly stated in the passage. This introduction consisted of sentences taken verbatim from the passage and following the original sequence. A textually implicit introduction provided explicit representation of information which was implied in the passage; thus, this type of introduction provided explicit linkages (e.g., connective words or phrases) between propositions when such linkages were merely implied in the passage. A scriptally implicit introduction presented substantively new, topic related information which was neither stated nor implied in the passage. Further, this type of introduction explicitly linked this new information to an idea in the passage. (See the Appendix for a sample stimulus passage and examples of introductions.) All introductions contained between nine and 11 propositions. For validation purposes, seven independent judges (graduate students in reading education) classified each unlabeled introduction as either textually explicit, textually implicit, scriptally implicit, or "other." Agreement among the individual judges and the investigator was 100 per cent.

Directions for free recall. In order to encourage all students to include in their recalls both text based and knowledge based information, the following written directions were provided for the free recall task: "Write down everything you can remember about what you read, and anything else it made you think about."

Probe questions. In order to give students an opportunity to add other information they might have remembered but had not included in free recall, six general, open-ended probe questions were presented immediately following the free recall task. Three questions addressed text-based information (e.g., "What else was there in the article that you did not write down before? What other facts or details from the article can you remember?"); three other questions addressed knowledge-based information (e.g., "What other things did the article make you think about? What else did you know about this topic?").

Procedure

All students participated in two group testing sessions within their home school; the second session was held a week following the first. In the first session, students were randomly assigned to one of four introductory conditions and to one of seven passages. Experimental subjects received either a textually explicit, textually implicit, or scriptally implicit introduction; control subjects received written directions to "Sit quietly for a minute or two while you think about what you do when you are reading something carefully to understand it." The students were informed that the investigator was interested in "how people learn from reading," that they would read a short article followed by "other activities," that they would not be graded or marked on their performance, and that they would be permitted to complete the booklets in a self-paced manner. During the first session, all subjects engaged in the following sequence of tasks: (1) reading a written introduction or directions to "think about . . . reading . . ." (2) reading one stimulus passage (3) engaging in a buffer activity (copying geometric designs) (4) writing a free recall (5) writing responses to probe questions, and (6) engaging in a filler activity (solving two pages of simple addition and subtraction problems). During the second testing session, students wrote another free recall, responded to probe questions, and engaged in a filler activity (writing responses to six, open-ended questions about their attitudes toward reading (cf., Canney and Winograd, 1979).

Scoring

The seven passages and the students' immediate and delayed recall protocols were segmented into propositions using Frederiksen's (1975) system of text analysis as modified by Tierney and his colleagues (Tierney and Bos, 1979; Tierney, Bridge and Cera, 1979). The propositions from each recall protocol were then compared to the propositions from the corresponding passage and classified according to a scoring system developed by Wixson (1979) and the present investigator. The scoring system utilized the categories in the Pearson and Johnson (1978) taxonomy of question-answer relations; propositions were classified as either textually explicit, textually implicit, scriptally implicit, text-based error, or script-based error.

Textually explicit propositions. A proposition was classified as textually explicit (TE) if its constituent concepts and/or relations were identical to or a paraphrase of those in a proposition derived from the passage. (Paraphrases were defined as synonomous or anaphoric referents for concepts within passage propositions.)

Example: Passage: The river god was angry.
Recall: He was mad.

Textually implicit proposition. A proposition was classified as textually implicit (TI) if it represented a proposition that was implied in the passage. Four types of text-based inferences were included within this category:

(1) concepts which were superordinates or subordinates of concepts within propositions from the passage.

Example: Passage: They (people) threw dolls and gifts into the river.
Recall: Little children threw dolls and things into the river.

(2) recall propositions which represented presuppositions or consequences inferrable from passage propositions.

Example: Passage: Egypt began to build the Aswan Dam in 1960.
Recall: Modern Egyptians solved the river problem.

(3) recall propositions which represented a synthesis of segments from separate propositions within the passage.

Example: Passage: For thousands of years the Nile River has made the land around it rich. Then there were bad floods.
Recall: The Nile River flooded the land around it.

(4) recall propositions which represented an explicit logical relation between propositions when that relation was implied in the passage.

Example: Passage: Then the crops died and people went hungry.
Recall: The people went hungry because there wasn't enough food.

Scriptally implicit proposition. A recall proposition was classified as scriptally implicit (SI) if it was plausible and if it represented new concepts or new information which were not stated or implied in the passage.

Example: Passage: (Indicated only that the Nile River was in Egypt.)
Recall: The Nile River is in a tropical region.

Implausible propositions. A recall proposition was classified as implausible if it represented either a text-based error or a script-based error. A text-based error referred to a recall proposition in which the implausible element was derived from concepts stated within the passage.

Example: Passage: Egypt began to build the Answan Dam in 1960.
Recall: They built the Nile River.

A script-based error referred to a proposition in which the implausible element was generated from the reader's own knowledge base rather than from the passage.

Example: Passage: But the ancient people of Egypt could not control the River.
Recall: The Indians could not control the river.

To establish inter-rater reliability for the scoring procedures, 10 recall protocols from the present study were analyzed and scored by a colleague who had utilized the same procedures in a related study, and 10 protocols from the colleague's study were analyzed and scored by the investigator. The investigator and second scorer had 97 per cent agreement on segmenting the recalls into propositions and 88 per cent agreement on classifying the individual propositions.

RESULTS

The criterion measures were the number of textually explicit propositions, textually implicit propositions, scriptally implicit propositions, and textually correct propositions generated in the immediate recall and in the delayed recall. (The textually correct measure was formed by adding together the number of textually explicit and textually implicit propositions.) For each student, the number of propositions of each type was expressed as a proportion of the total number of propositions produced in either the immediate or delayed recall. Subsequent data analyses were performed on the 102 students who had produced a minimum total of five propositions in their immediate recalls and in their delayed recalls.

First, several one-way analysis of variance procedures, with school or passage as the separate between subjects factor, were performed on each recall measure. Results indicated that there were no reliable differences due to school or passage. Therefore, the data were combined for further analyses.

Next, a 4 (introductory condition) x 2 (time of recall) analysis of variance with repeated measures on time was carried out on each recall measure. (See Table 1 for the mean amounts of the different recall measures.) Results of the analysis on the proportion of textually explicit propositions indicated that there was not a significant main effect for introductory condition, $F(3,98) = 1.05, p < .37$, but that there was a significant main effect for time, $F(1,98) = 19.09, p < .0001$. Thus, all subjects, regardless of the type of introduction they received, produced significantly greater proportions of textually explicit propositions in immediate recall than in the delayed recall. With respect to the textually implicit measure, there was neither a significant main effect for introductory conditions, $F(3,98) < 1.00$, nor a significant main effect for time, $F(1,98) < 1.00$. For the scriptally implicit measure, there was a significant main effect for introductory condition, $F(3,98) = 6.60, p < .0005$, and a significant effect for time, $F(1,98) = 15.63, p < .0001$. Thus, all subjects generated significantly greater proportions of scriptally implicit propositions in their delayed recalls than in their immediate recalls. The results of the analysis on the textually correct measure indicated that the effect for introductory condition approached significance, $F(3,98) = 2.40, p < .07$, and that the effect for time was significant, $F(1,98) = 18.49, p < .0001$. Thus, all subjects generated greater proportions of textually correct information in the immediate recall than in the delayed recall. The introductory condition × time of recall interaction was not significant for any of the recall measures.

Table 1
Mean Amounts[a] of Recalled Propositions
by Introduction Group

	Immediately Recalled Measures				**Delayed Recalled Measures**			
	TE	**TI**	**SI**	**TEXCOR**	**TE**	**TI**	**SI**	**TEXCOR**
TE Group (n = 24)	.36	.42	.12	.78	.30	.40	.18	.70
TI Group (n = 28)	.38	.43	.13	.82	.26	.41	.25	.67
SI Group (n = 27)	.32	.38	.25	.70	.26	.39	.31	.65
Control Group (n = 23)	.29	.37	.18	.66	.26	.37	.20	.63

[a] Proportions of the total number of recalled propositions.

Separate one-way analysis of variance procedures, with introductory condition as the between subjects factor, and subsequent post hoc analyses using Scheffé's techniques (.05 level) were additionally performed on the immediate and delayed scriptally implicit measures and on the immediate and delayed textually correct measures. Results of the analysis on the immediate scriptally implicit measure indicated a reliable effect for introductory condition, F (3,98) = 6.49, $p < .0005$. Pairwise comparisons revealed that students who had received the scriptally implicit introductions generated significantly greater proportions of scriptally implicit propositions (M = .25) than students who had received either the textually explicit introductions (M = .12) or the textually implicit introductions (M = .13). A significant F ratio was also found on the delayed scriptally implicit measure, F (3,98) = 3.52, $p < .02$. Pairwise comparisons indicated that there was a significant difference between the group given the scriptally implicit introductions (M = .31) and the group given the textually explicit introductions (M = .18). Results of the analysis on the immediate textually correct measure indicated a significant F ratio, F (3,98) = 4.50, $p < .006$; pairwise comparisons indicated that there was a significant difference between the group given the textually implicit introductions (M = .82) and the control group (M = .66). Analysis of the delayed textually correct data revealed that there was not a significant effect for introductory condition, F (3,98) < 1.00, $p > .05$.

DISCUSSION

All students, regardless of the type of introductory condition received before reading a short, informational passage, recalled significantly smaller proportions of textually explicit information and significantly smaller proportions of textually correct information after an interval of a week than immediately after reading. Additionally, all students produced greater proportions of scriptally implicit information in their delayed recalls than in their immediate recalls. These findings are consistent with results reported by Dooling and Christiaansen (1977) and Sulin and Dooling (1974) in support of the constructivist view of memory for prose (Bartlett, 1932). The constructivist view holds that with the passage of time subjects forget specific, textual details, and at the time of recall, construct details on the basis of their current knowledge about the general theme.

However, all students, regardless of introduction type or no introduction, produced similar amounts of textually explicit information in both the immediate and delayed recall situations. These results may be due to the reading proficiency of the subjects and the characteristics of the passages used in the study. Inasmuch as the subjects were fifth grade students with average and above average reading skills and the passages were written at the fourth grade level of reading difficulty, it is not surprising that these skilled readers would retain a basic corpus of explicitly stated text information. The present findings of no significant differences on the relative amounts of textually explicit information among the different introduction groups may be consistent with results emerging from Swaby's (1978) study which investigated the effects of an advance organizer, a vocabulary introduction, and no prereading treatment on good and poor sixth grade readers' comprehension of text. Swaby found that the comprehension of the good readers was not reliably affected by the type of introductory treatment; the poor reade[illegible]er study, however, performed better following the vocabulary treatment than [illegible] other treatments.

In the present study, the introductions did not significantly affect subjects' recall of textually implicit information. These results suggest support for Paris and Upton's (1976) findings that fifth grade children are capable of inferring a variety of implied relations; the results are also consistent with Kintsch's (1974) contention that text-based inferences are necessary for the construction of a coherent representation of a text. Thus, it appears that children in the fifth grade, like Kintsch's adult subjects, spontaneously generate text-based inferences in the process of formulating a coherent representation of a text.

A further finding from the present study indicates that students who received the scriptally implicit introductions generated significantly greater amounts of scriptally implicit information than students who received the textually explicit introductions. Both groups, however, as well as the other treatment group and the control group, generated similar amounts of textually explicit information. Since scriptal information is generated by readers on the basis of their own world knowledge (Pearson and Johnson, 1978), rather than derived from the text, the presence of scriptal information in readers' recalls is viewed as evidence of the integration of new, text information with old, existing knowledge (Spiro, 1977). Within such a view, the results of the present study suggest that scriptally implicit introductions promote greater integration of text information with readers' prior knowledge, whereas textually explicit introductions appear to promote such integration to a lesser extent.

REFERENCES

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 1960, *51*, 267-272.

Bartlett, F. C. *Remembering.* Cambridge, England: The University Press, 1932.

Bransford, J. D. and Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 717-726.

Bransford, J. D. and McCarrell, N. S. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weiman and S. D. Palermo (Eds.), *Cognition and the symbolic processes.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.

Brown, A. L. and Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, 1977, *48*, 1-8.

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., and Lawton, S. C. Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development*, 1977, *48*, 1454-1466.

Canney, G. and Winograd, P. *Schemata for reading and reading comprehension.* (Technical Report No. 120). Urbana, Ill.: University of Illinois, Center for the Study of Reading, April, 1979.

Dooling, D. J. and Christiaansen, R. E. Levels of encoding and retention of prose. In G. H. Bower (Ed.), *Psychology of learning and motivation.* New York: Academic Pross, 1977, 5-39.

Dooling, D. J. and Lachman, R. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, *88*, 216-222.

Dooling, D. J. and Mullet, R. L. Locus of effects in retention of prose. *Journal of Experimental Psychology*, 1973, *97*, 404-406.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458.

Gagne, E. D. Long term retention of information following learning of text. *Review of Educational Research*, 1978, *48*, 629-665.

Holt Basic Readers. New York: Holt, Rinehart and Winston, 1977.

Keys to Reading. Oklahoma City: Economy Co., 1975.

Kintsch, W. *The representation of meaning in memory.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.

Liddle, William. *Reading for concepts* (Book C). St. Louis, Mo.: McGraw-Hill, 1977.

Mayer, R. E. Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science*, 1979, *8*, 133-167.

Paris, S. G. and Upton, L. R. Children's memory for inferential relationships in prose. *Child Development*, 1976, *47*, 660-668.

Pearson, P. D. and Johnson, D. D. *Teaching reading comprehension*. New York: Holt, Rinehart & Winston, 1978.

Pichert, J. W. and Anderson, R. C. Taking different perspectives on a story. *Journal of Educational Psychology*, 1977, *69*, 309-315.

Reading Basics Plus. New York: Harper and Row, 1976.

Schallert, D. L. Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 621-632.

Shimmerlik, S. M. Organization theory and memory for prose: A review of the literature. *Review of Educational Research*, 1978, *48*, 103-120.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro and W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Sulin, R. A. and Dooling, D. J. Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology*, 1974, *103*, 255-262.

Swaby, B. E. R. The effects of advance organizers and vocabulary introduction on the reading comprehension of sixth grade students. (Doctoral dissertation, University of Minnesota, 1977.) *Dissertation Abstracts International*, 1978, *39*, 115A-116A.

Tierney, R. J. and Bos, C. S. Personal communication, July 1979.

Tierney, R. J., Bridge, C. and Cera, M. J. The discourse processing operations of children. *Reading Research Quarterly*, 1978-79, *14*, 539-573.

Wixson, K. K. *The effects of postreading questions on children's discourse comprehension and knowledge acquisition*. Unpublished doctoral dissertation, Syracuse University, 1980.

APPENDIX

SAMPLE INTRODUCTIONS TO THE NILE RIVER PASSAGE

TE Introduction

For thousands of years, the Nile River has made the land around it rich. But, the ancient people of Egypt could not control the huge river. Often the river overflowed too much. At other times, there was too little water from the Nile. The people said these things happened because the river god was angry. Egypt began to build the Aswan Dam in 1960.

TI Introduction

Although the Nile river made the land around it rich, it caused problems for the ancient Egyptians. When the huge river overflowed too much, it caused bad floods. At other times the crops died because the river did not overflow enough. In 1960, the people of Egypt began to build a dam for the river.

SI Introduction

In ancient times many people believed in different kinds of gods. They believed that the gods controlled the forces of nature, such as the wind, the sun, and water. The ancient Egyptians thought that a river god made the Nile River overflow. When the river overflowed, it watered the farms and left fertile mud on the fields.

SAMPLE STIMULUS PASSAGE

The Nile River

For thousands of years the Nile River has made the land around it rich. But the ancient people of Egypt could not control the huge river. Often the river overflowed too much. Then there were bad floods. At other times there was too little water from the Nile. Then the crops died, and the people went hungry. The people said these things happened because the river god was angry.

They threw dolls and gifts into the river. They hoped that these gifts would make the river god happy. Today, we know that wet seasons can make a river overflow. Dry seasons can make a river dry up.

A dam can stop the flow of a river. A dam can store water in wet seasons for use in dry seasons. Egypt began to build the Aswan Dam in 1960. This dam is 3 miles long and 350 feet high. It can control floods on the Nile River.

The Egyptians of today do not need to try to keep a river god happy.

© 1977, 1970 by McGraw-Hill, Inc. Used by permission of the publisher.

FRANK SPENCER
University of Arizona

MICHAEL JOHNSTON
Tuscon Unified School District

WILBUR AMES
University of Arizona

THE EFFECT OF MANIPULATING THE ADVANCE ORGANIZER AND OTHER PREREADING STRATEGIES ON COMPREHENSION OF ABSTRACT TEXT

Two decades of research on many aspects of the advance organizer (AO) have produced conflicting results. Reviews of the literature have indicated that the AO can have some positive effect on learning from text under certain conditions, as when subjects are of lower verbal or reading ability, or when the material is unfamiliar or poorly organized (Mayer, 1979). Research on the use of the AO with abstract material has been limited, perhaps due to Ausubel's statement (1968) that abstract material has inherent organizing properties, such that the AO would not aid the learning process when used with such material. Since experience indicates that many high school students often have difficulty learning from abstract material, it seems possible that any inherent organization of the concepts in abstract material may not be perceived nor utilized by such students.

In a previous study (Johnston and Spencer, 1979), attempts were made to manipulate the AO for use with abstract material. That work has been extended in this study. It was hypothesized that an AO written at a lower level of abstraction, using conceptual analogies familiar to the students, would alert students to those conceptual organizers inherent in abstract material, or provide such organization if it was in fact *not* inherent in the material.

A second hypothesis in this study concerned the use of a prereading strategy which points out to the reader the surface organization of the text. Meyer (1977) suggests that utilizing the author's organization appears to be the most efficient strategy for most school learning situations. Such organization guides apply most directly to expository materials (Herber, 1978), and appear to be especially appropriate for lower ability students when reading for factual information (Hartley and Davies, 1976). It has been suggested, however, that imposing an outline on material that is already highly organized may interfere with processing (Northrop, 1952; Jenkins and Bausell, 1976).

A tentative conclusion to be drawn from a review of the literature suggests, then, that the conceptual organization inherent in abstract text and the surface organization of expository material are sufficient for understanding and retaining the information, both higher-order concepts and factual details, presented in such texts. Imposing prereading organizers supposedly risks interfering with the positive effects of these aspects of such texts on the reader: too much of a good thing is not good.

In the current study, an AO and a text organizer (TO) were each used as a prereading strategy for enhancing learning from abstract, organized material. In addition, a third treatment consisted of a combined AO and TO, to investigate the possibility that addressing the conceptual *and* surface organizations of a text in a prereading activity would lead to better comprehension and retention.

PROCEDURE

Subjects were sophomores in English classes in a public high school in Tucson, Arizona. Since this study proposed to look at only those students with average reading ability, only the data from those with stanine scores of 4 through 6 on the California Achievement Test (administered to them as freshmen the previous Spring) were analyzed from the total class groups. The research instrument was presented as a study skills survey, the results of which would be used as a basis for establishing a study skills program.

The target instructional material, chosen from a high school government textbook (McClenaghan, 1978), described four theories of the origin of the state. The first treatment (an advance organizer using conceptual analogies) consisted of an analogy to the family, in which sources of authority in the family which parallel the four theories of the origin of the state were discussed. The second treatment gave students an overview of the organizational aspects of the target material. The third treatment was a combination of the first two. The control group received a buffer passage which dealt with the development of the Constitution. All treatments were approximately 300-350 words with a Fry readability range of 9th-10th grades.

Instrumentation used included a pretest of prior knowledge, in which five multiple-choice questions were buried in a 20-item test. The five questions covered the key concepts that were presented in the target material. The posttest consisted of 11 factual and 6 higher-order questions, and was used for both immediate and delayed posttesting. It was shown to have a reliability coefficient of .62 using the split-half method with correction.

Three hundred eighty-four pretested students were randomly assigned to each of the three treatment groups and the control group. One week after the pretest, the treatments and immediate posttest were administered by the experimenter to 378 students, 199 of whom fell into the average reading ability and low prior knowledge categories. After the delayed posttest was given three weeks later, data of 124 subjects were analyzed by ANOVA.

RESULTS

Immediate and delayed posttest results were first analyzed for total correct scores. Subsequently, results for factual and higher-order questions were analyzed. Table 1 presents the raw score data.

TABLE 1

Mean Raw Scores of Treatment Groups on Immediate and Delayed Posttests

	Immediate Posttest				*Delayed Posttest*			
Treatment	*n*	Total	Factual	Higher Order	*n*[*]	Total	Factual	Higher Order
AO	55	7.45	5.31	2.14	31	5.06	3.35	1.71
TO	47	8.49	6.11	2.38	31	5.39	3.52	1.87
AO + TO	47	6.92	4.81	2.11	31	4.97	3.32	1.65
Control	50	6.20	4.40	1.80	31	4.77	2.94	1.84

[*] Cell sizes equalized by random deletion.

Significant differences were found between treatment groups on the immediate posttest for total scores, *F* (3,195) = 6.52, $p \leq .01$, and for factual scores, *F* (3,195) = 7.12, $p \leq .01$. Post hoc analyses by Scheffé method indicated that the Text Organizer group scored significantly higher than both the AO + TO group and the control group, for both total and factual-only scores. No significant differences were found for higher-order questions. The delayed posttest showed no significant differences among treatment groups, although the total and factual means maintained the same ranking as on the immediate posttest.

DISCUSSION

The text organizer treatment was an attempt to enhance the readers' interaction with a text which was well organized on the surface level. This prereading strategy appears to be effective in learning factual material if the imposed organization mirrors that in the text.

The advance organizer treatment attempted to manipulate the conceptual organization of the reader's cognitive structure in order to enhance the learning of new concepts from abstractly written material. By using language and concepts familiar to the learner and concrete analogies, the AO treatment was to capitalize on the organizing properties supposedly inherent in abstract text. That the AO treatment produced good (but not significantly better) results indicates that any interference of this prereading strategy was minimal. That positive facilitation of learning is not apparent in the results may be due to one of several possibilities. It could be that all pieces of abstractly worded text do *not* have built-in organizers, in which case an AO would affect learning from that text differentially. If abstract text does have inherent conceptual organizers, the efficacy of an imposed AO may depend on its goodness of fit to those inherent organizational properties of the text. It seems reasonable that the surface text organization was accurately reflected in the TO treatment, but the deep structure of the text may not have been similarly mirrored by the AO.

The combination AO + TO treatment was meant to capitalize on the learning-facilitating properties of each individual strategy. That it did not perform better than either individual treatment suggests the possibility that some overload may occur in the use of prereading stragegies. The combined approach attempted to direct the reader's attention to both the conceptual and text organizations, and it may be that students had trouble utilizing these separate approaches while interacting with the text.

In general, the results indicate that those groups with some preparation for reading performed better than those with no preparation for reading. That the significant results were not maintained through the delayed posttest is not greatly surprising. The factual responses which apparently accounted for the significant differences on the immediate posttest would be much more subject to forgetting, especially without reinforcement, than learning that may have taken place on the conceptual level.

CONCLUSIONS

This investigation was primarily concerned with the effect of two different prereading strategies on high school students' comprehension of abstract, organized material. Based on the results, and under the conditions of this study, it can be concluded that providing the learner with an overview of text organizational characteristics can significantly enhance the processing of abstract, unfamiliar and relatively well-organized text. This appears to lend support to Meyer's (1977) contention that this prereading strategy is perhaps the most efficient one available to the classroom teacher.

Although the results of the advance organizer group were not significantly higher than the control group, the resulting trend toward better comprehension does indicate that perhaps under different circumstances (e.g., abstract material that is not so well organized) AO's might significantly improve learners' comprehension. If the inherent conceptual organization depends on well organized surface structure, then authors of abstract material should ensure that their abstract concepts and critical propositions are logically developed and organized.

This study neither proves nor disproves the viability of using an advance organizer with abstract, unfamiliar material. If use of the text organizer resulted in significant differences because it closely followed the surface organization of the text, then it is possible that the advance organizer *could* perform similarly well if it closely paralleled the conceptual organization inherent in abstract text. Further research is called for in order to determine the nature of such inherent organization and how it can be utilized to enable the production of efficacious prereading activities. As it stands now, the amount of time necessary to develop a beneficial advance organizer for abstract, unfamiliar and unorganized text might be too impractical and inefficient for the classroom teacher.

REFERENCES

Ausubel, D. *Educational psychology: A cognitive view.* New York: Holt, Rinehart & Winston, 1968.

Downie, N. W., & Heath, R. W. *Basic statistical methods* (2nd ed.). New York: Harper & Row, 1965.

Hartley, J., & Davies, I. Preinstructional strategies: The role of pretests, behavioral objectives, overviews and advance organizers. *Review of Educational Research,* Spring, 1976, *46,* 239-265.

Jenkins, J. R., & Bausell, R. B. Cognitive structure variables in prose learning. *Journal of Reading Behavior,* 1976, *8* (1), 47-66.

Johnston, M. & Spencer, F. *The effects of manipulating the advance organizer on comprehension of abstract material.* Paper presented at the annual meeting of the National Reading Conference, San Antonio, December, 1979.

Mayer, R. E. Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science,* 1979, *8,* 133-167.

McClenaghan, W. *Magruder's American Government.* New York: Allyn & Bacon, 1978, 18-19.

Meyer, B. Organization in prose and memory: Research with application to reading comprehension. In P. David Pearson (Ed.), *Reading: Theory, research and practice,* 26th Yearbook of the National Reading Conference, Clemson, SC: National Reading Conference, 1977.

Northrop, D. S. *Effects on learning of the preeminence of organizational outline in instructional films.* (Instructional Film Research Program, Pennsylvania State University) Port Washington, NY USN Training Device Center, 1952. (ERIC Document Reproduction Service No. ED 002 701)

TINA JACOBOWITZ
Montclair State College

THE EFFECTS OF MODIFIED SKIMMING ON COLLEGE STUDENTS' RECALL AND RECOGNITION OF EXPOSITORY TEXT

Current researchers in reading recognize that advanced knowledge of text macrostructure, that is the general scheme or gist of a passage, aids comprehension (Frederiksen, 1975; Kintsch, 1974; Meyer, 1975). and recommend such strategies as structured overviews (Herber, 1978) and pre-questioning techniques (Frase, 1975) to facilitate understanding of text. These teacher-guided, pre-reading strategies are often designed to provide the reader with an overview of the text, which serves as a framework into which the reader can fit new, incoming information.

Many college students have limited understanding of text structure and organization. In addition, there is evidence that the slow and inefficient word coding processes of less skilled readers result in loss of memory of the gist or theme of the text derived from previous sentences (Perfetti & Lesgold, 1977). Such research indicates the value of providing college students, particularly those who are not skilled readers, with an understanding of the macrostructure of the text prior to reading it. While the strategies mentioned previously serve this purpose, they rely heavily on teacher/student interaction. In that the ultimate goal of reading instruction is to enable students to learn on their own, it is crucial that they learn techniques which will facilitate independent acquisition of text macrostructure.

There is some research which suggests that previewing, which entails prereading particular portions of text, may be an effective strategy for gaining insight into the general scheme of the text. For example, Anderson (1978) reports that the skilled readers he observed previewed the text in some manner in order to get an overall picture of text organization. One previewing technique, skimming, provides additional insight into text macrostructure by exposing the reader to many of the major points intended by the author.

Macrostructure may be viewed as a higher-order unit of organization which summarizes the main points of the text (Waters, 1980). There is some evidence that prior knowledge of main ideas increases recall of important information (Rickards, 1975). Adapting the concept of advance organizers espoused by Ausabel (1960), Rickards interspersed main idea sentences in the text, that is, one main idea sentence was inserted before every two paragraphs in the text. These superordinate prestatements yielded more recall of superordinate concepts than any other condition, a finding which seems to be compatible with results of discourse analysis studies.

The present study was designed to determine the effects of using a modified or mock skimming technique on the reading comprehension of college students. According to Brown (1975), the main ideas of sixty to ninety percent of expository paragraphs are contained in the first sentences of these paragraphs, with the last sentence being the next most likely place to find them. When using the skimming strategy, the first and last paragraphs are read in their entirety, and the first and possibly the last sentences, of the intervening paragraphs are read. Thus, skimming exposes the reader to many of the main ideas which comprise the macrostructure of the selection. In this study, a modified skimming technique was developed which enabled students to preview the selection in a manner somewhat similar to actual skimming, yet ensured that students would not read more than those portions of the text specified above. This procedure involved providing students with a preview sheet which contained the first and last paragraphs and the first sentences of the remaining paragraphs, of the text.

METHOD

Sample

The 129 students involved in this study were undergraduates enrolled in 13 college reading improvement and teacher-training courses at Montclair State College. Participants were assigned to one of three groups such that an approximately equal number of students from the two courses were in each group.

Materials

To determine students' reading abilities, the Directed Reading Test (Committee on Diagnostic Reading Tests, 1966), a standardized college reading test was administered to all subjects two months previous to the experiment. In addition, all participants took a 12-item multiple choice preknowledge assessment test three weeks prior to the experiment, the scores of which were used as a covariate in the data analysis. This test consisted of 36 multiple-choice questions. Twelve of the items related to the major ideas in the experimental selection, while the remaining 24 questions dealt with two totally unrelated topics and served as distractors.

The stimulus selection, consisting of 13 paragraphs, was a modified version of a story entitled "Why Birds Sing" (Peterson, 1961). This article was selected for several reasons: (a) it was felt that while students would have some familiarity with the subject, they would not have sufficient background knowledge to answer the questions without reading the text. This was substantiated by the low scores achieved on the preknowledge test. (b) The selection was written in expository style similar to textbook writing. (c) While most of the main ideas were found in the beginnings of the paragraphs, there were some paragraphs with unstated main ideas, main ideas located in the middle or at the end of paragraphs, and main ideas made up of more than one of the sentences in the paragraph. Thus, the experimental group did not simply receive a list of main ideas on the preview sheet. (d) Prior to the experiment, the selection was administered to non-participating low-reading ability students who had little difficulty reading it. Thus, the selection was deemed suitable for use in this study.

Three experts in reading were asked to outline the selection using the overriding topic "The Nature and Functions of Bird Songs" as a guideline. They were asked to indicate the superordinate or main ideas and the subordinate ideas or details in their outlines. In that there was high agreement among the three individuals, the outlines were consolidated and the resulting outline served as a model for scoring the protocols. In addition, a 20-item multiple choice recognition test consisting of 12 main idea and 8 detail questions was developed. The three individuals categorized the items as either main idea or detail questions; there was 100% agreement on the classification of each item.

Procedures

There was one experimental group and two control groups in this study. All data were collected during the regularly scheduled meetings of each class. Participants received packets of materials containing the selection and the 20-item recognition test. In addition, the experimental, or Skim Group, received the preview sheet which they were instructed to read prior to reading the entire selection. Control Group 1 did not receive the preview sheet; this group was asked to read the selection once. To account for the fact that the Skim Group read some portions of the material two times, students in Control Group 2 were instructed to read the entire selection twice. All students were asked to record time elapsed during their first reading of the entire selection. Following the reading, students were instructed to write as many ideas as they could recall from the selection, and then to respond to the 20 questions.

Data Analysis

The general design for this study involved Reading Technique (modified skimming, one reading, two readings) as an independent variable and six comprehension measures as dependent variables. Measures of comprehension included total scores on the recognition and recall tasks, as well as main idea and detail subscores for these two tasks. Data were analyzed by using analysis of covariance and follow-up pair-wise comparisons using the Tukey HSD technique. The covariates were the Diagnostic Reading and preknowledge assessment test scores, and the amount of time, in seconds, taken by each student for the first reading of the entire selection.

RESULTS AND DISCUSSION

A random sample of 20 recall protocols was scored for number of main ideas and details by two individuals who matched students' ideas to those on the model outline. Scorers were blind to the assigned treatment for each group. This procedure produced interrater reliability coefficients of above .98 for main ideas, details, and total number of ideas recalled.

The results of the analysis of covariance tests indicated significant differences among the adjusted means on all the dependent variables. Specifically, significant differences were found for total recognition scores ($F\ (2, 123) = 6.85$, $p < .002$), recognition of main idea scores ($F\ (2, 123) = 4.60$, $p < .01$), recognition of detail scores ($F\ (2, 123) = 4.32$, $p < .02$), as well as for total number of ideas recalled ($F\ (2, 123) = 3.91$, $p < .02$), number of main ideas recalled ($F\ (2, 123) = 10.08$, $p < .001$), and number of details recalled ($F\ (2, 123) = 10.04$, $p < .001$). Follow-up pair-wise comparisons showed that students who mock-skimmed the selection prior to reading it achieved significantly higher scores on the recognition test than did students in Control Group 1, who read the selection once ($p < .05$). This finding is similar to that of Reder & Anderson (1980) who found that individuals who read summaries of text demonstrated better comprehension than students who read the entire text. While the preview sheet provided to subjects in the Skim Group was not a summary, it did contain many of the main ideas of the selection due to its expository nature. In addition, students in the Skim Group performed similarly on the recognition task to students in Control Group 2, who read the entire selection twice. Both of these findings were the same when scores were analyzed according to main idea and detail subscores.

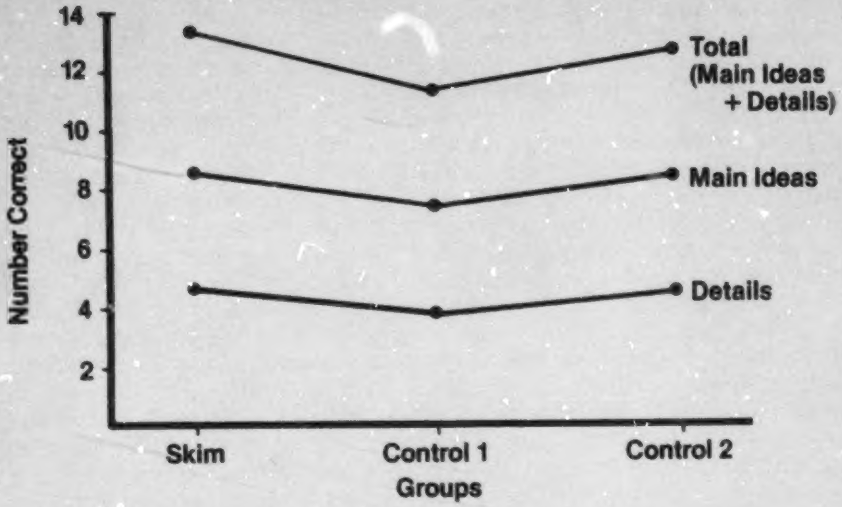
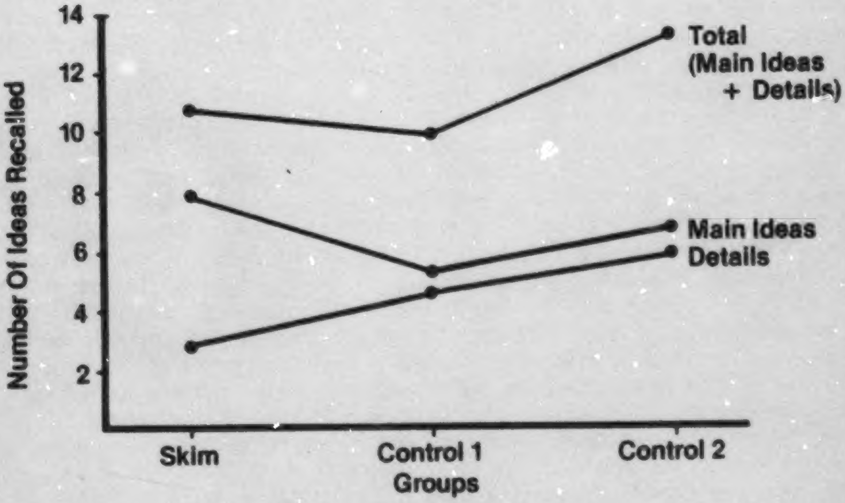


Figure 1. Mean scores on the recognition task

Figure 2. Mean scores on the recall task



Results were somewhat different for the recall aspect of the experiment. As on the recognition test, the Skim Group achieved total recall scores similar to Control Group 2. However, for the recall, only Control Group 2 recorded a significantly greater number of ideas than did Control Group 1 ($p<.05$). This is due to the fact that students in the Skim Group recorded significantly fewer details than did students in the two control groups. Students in the Skim Group did however, record significantly more main ideas than those in Control Group 1, and about the same number as those in Control Group 2. That the Skim Group did not include as many details in their recall protocols does not necessarily mean that they remembered less than the others; in fact, their superior performance relative to Control Group 1 on the recognition of details subtest lends further support to the notion that the modified skimming provided a framework into which students could integrate additional information. Analysis of the protocols of students in the Skim Group suggests that these individuals were more aware of overall text organization and chose to include only those ideas which were most relevant to the gist of the selection than did students in the two control groups. Protocols of students in the Skim Group consisted primarily of main ideas and were more like summaries of the text than were those of the control groups. Students in these groups wrote down as many ideas as they could remember, apparently paying less attention to the organization and important points of the selection.

Discussion following the experiment provided further insight into the benefits of modified skimming prior to reading a selection. Students in the experimental group felt that the mock skimming focused them on the organization and main ideas of the article before reading it, while those in Control Group 2 said that the second reading helped them to remember more, but that the first reading did not focus them on the overall organization or major points of the selection.

CONCLUSIONS AND IMPLICATIONS

The findings of this study suggest that reading information-rich portions of expository text prior to reading the entire selection is an effective and efficient previewing strategy. Further investigation is needed, however, to determine if the results of this study, in which a contrived procedure was employed for experimental purposes, can be replicated using a more natural skimming strategy and various text selections. Students should be taught to preview text independently using a skimming strategy, without the aid of a preview sheet. It is possible that these students will exhibit even better comprehension than students who read a selection twice because they would be taking a more active role in the process. In this study, students were not told that the preview sheets contained many of the main ideas of the selection. A more active skimming procedure would enable a reader to actively search for the main points rather than to just passively read the first sentences of the paragraphs. Further, it would encourage students to generate hypotheses about what is contained in the selection which they could then confirm or disconfirm when reading the entire selection.

Results of this study also suggest that previewing text using a skimming strategy may be more effective for recall of main ideas than of details. Thus, students should be aware of their purpose for reading before opting to use this particular technique.

In summary, the findings of this investigation imply that skimming a selection before reading it may provide students with insight into the overall organization of the text, as well as expose them to many of the superordinate ideas. Skimming, therefore, is one previewing technique which may enhance comprehension, particularly of main ideas, by enabling the reader to grasp independently the macrostructure of the text.

REFERENCES

Anderson, T. H. *Study skills and learning strategies*. (Technical Report No. 104). Urbana: University of Illinois, Center for the Study of Reading, September, 1978.

Ausabel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 1960, *51*, 267-272.

Brown, J. I. *Reading power*. Lexington, Mass.: D.C. Heath & Co., 1975.

Committee on Diagnostic Reading Tests, *Diagnostic Reading Tests*, Mountain Home, North Carolina: Committee on Diagnostic Reading Tests, Inc., 1966.

Frase, L. T. Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 1975, *87*, 628-635.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 317-458.

Herber, H. L. *Teaching reading in content areas*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1978.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1974.

Meyer, B. Identification of the structure of prose and its implication for the study of reading and memory. *Journal of Reading Behavior*, 1975, *7*, 7-47.

Perfetti, C. & Lesgold, A. Discourse comprehension and sources of individual differences. In M. Just & P. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Lawrence Ehrlbaum Associates, 1977.

Peterson, R. T. Why birds sing. *The New York Times Magazine*. October 22, 1961.

Reder, L. M. and Anderson, J. R. A comparison of texts and their summaries: memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 121-134.

Rickards, J. P. Processing effects of advance organizers interspersed in text. *Reading Research Quarterly*, 1975-1976, *11*, 599-622.

Waters, H. S. Superordinate-subordinate structure in semantic memory: the roles of comprehension and retrieval processes. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 587-597.

BONNIE C. GREENSLADE
Northern Illinois University

THE QUALITY, QUANTITY, AND VARIETY OF TEACHER AND PUPIL QUESTIONS AND PUPIL RESPONSES IN DISCUSSIONS FOLLOWING GROUP MAPPING ACTIVITIES

The need for instructional practices which encourage and develop comprehension has long been recognized. Taba stated the need for giving systematic thought to the ways in which knowledge acquisition also becomes a method of inquiry and thought (1967, p. 179). Smith emphasized the crucial role of the teacher in developing comprehension through guiding and encouraging students to both invent and answer questions (1969, p. 5). However, Durkin (1978) in her investigation of reading comprehension instruction, found very little evidence of comprehension instruction taking place. She noted that attention directed to comprehension was chiefly through teacher questions aimed at assessment.

The nature of questions asked by teachers and the corresponding responses made by pupils has been examined by Wolf, Huck, & King (1967) and Davidson (1978) and a relationship was observed between the kinds of questions teachers asked and the level of critical responses from pupils. In both studies specific fact questions elicited non-critical responses while interpreting, inferring questions elicited critical levels of response from pupils. In light of these findings, it would seem appropriate for teachers to ask interpreting, inferring questions to stimulate critical level responses, representing higher levels of thinking, to maximize student comprehension of text.

Guszak (1967) found that the majority of questions asked by teachers required only literal comprehension. Hare and Pulliam (1980) in their recent study found that teachers have not significantly changed their questioning habits in the decade since Guszak's study.

Stauffer (1980) discusses the important role of students' questions and responses as well as student to student interaction in the comprehension of text. He stresses the art of asking relevant questions as one of the intellectual skills necessary to critical reading. He notes also the importance of student to student interaction in the development of critical thinking. He states that "scholars do not raise questions in an intellectual vacuum, rather, by virtue of the knowledge and experience available to them at the time they conjecture about answers. If they accept the questions that someone else raises, it is of even greater significance that they speculate about answers. By doing so they become personally committed and cognitively involved" (p. 185).

A procedure called the Group Mapping Activity (Davidson & Bayliss, 1978) provides for dynamic interaction through discussion after mapping. Teachers and pupils are actively involved in initiating questions and generating responses in an open communication network while maps are explained, discussed, and reacted to.

This study was designed to analyze questions and responses during the discussions following Group Mapping Activities to determine whether questions generated stimulated critical levels of response. In an informal analysis of questions asked by group members during a Group Mapping Activity, Davidson & Bayliss (1978) suggested the question types tended to be chiefly interpretive and evaluative. They noted that questions asked stimulated higher levels of thinking, based on the pupils' critical responses. The purpose of this exploratory study is to analyze pupil and teacher questions and pupil responses to formally examine the levels of thinking represented in discussions following Group Mapping Activities.

METHOD

Subjects who participated in this study were twenty-one eighth grade members of a language arts class in an Illinois rural community school district middle school. The group included all eighth grade pupils in the school district, representing various levels of ability.

Two short stories, "The Samaritan" by Richard Harper and "All the Years of Her Life" by Morley Callaghan were presented to all students by means of the Directed Reading-Thinking Activity (Stauffer, 1969). After pupils completed reading the passage they were instructed to retell the passage in writing. The Group Mapping Activity was explained to the pupils and they were instructed to map their perceptions of the passage. After mapping, pupils were encouraged to share their reasons for mapping as they did and for the inclusion of particular elements in their maps.

Two teachers participated in the study. Both had prior experience with Group Mapping Activities and with the Directed Reading-Thinking Activity. One teacher presented "The Samaritan" to the group by means of a Directed Reading-Thinking Activity, asked the pupils to map their perceptions of the story, and acted as facilitator in the discussion which followed the mapping. Subjects participating in the study had no prior experience with Group Mapping Activities or with Directed Reading-Thinking Activities. The passages read were, based on teacher judgment, at the independent or instructional reading level for all pupils participating in the study. There was a time interval of one week between the first session in which "The Samaritan" was read and mapped and the second session in which "All the Years of Her Life" was read and mapped. The map sharing time, the final fifteen minutes of each session, was recorded on a tape recorder. The investigator coded pupil and teacher questions and pupil responses according to the scales developed and used in the 1967 Ohio State University studies in critical thinking (Wolf, Huck & King).

SUMMARY OF RESULTS

An analysis of the questions generated during the discussion following the Group Mapping Activity after reading "The Samaritan" reveals that the teacher generated 5(26.3%) of total questions asked while pupils generated 14(73.7%). An analysis of the questions generated during the discussion following the Group Mapping Activity after reading "All the Years of Her Life" reveals that the teacher generated 14(50%) of total questions asked while pupils generated 14(50%). A combined analysis of the passages reveals that pupils generated 28(59.6%) of all questions asked. Of all questions asked by teachers and pupils 28(59.6%) were interpreting, inferring. Other questions asked by teachers and pupils were: evaluating, 5(10.6%); clarifying, 5(10.6%); gathering specific facts, 4(8.5%) analyzing, 3(6.3%); and applying, 2(4.2%). Of all teacher generated questions in the combined passages 10(52.6%) were interpreting, inferring. Other teacher generated questions were: gathering specific facts, 3(15.8%); clarifying, 2(10.5%); analyzing, 1(5.3%); applying, 1(5.3%); evaluating, 2(10.5%). Of all pupil generated questions in the combined passages, 18(64.3%) were interpreting, inferring. Other pupil generated questions were: gathering specific facts, 1(3.6%); clarifying, 3(10.7%); analyzing, 2(7.1%); applying, 1(3.6%); and evaluating, 3(10.7%).

A review of the combined pupil responses to questions generated during both discussions following the Group Mapping Activities reveals that 44(88.4%) of responses were at critical levels. There were 20(37.1%) critical level responses involving evaluating, using criteria, and/or supporting a stand. There were 16(29.6%) critical level responses involving inferring, applying, and/or illustrating. Other critical level responses involved theorizing and/or hypothesizing, 8(14.7%). Non-critical level responses were those involving literal memory, 6(11.1%), and guessing or random response, 4(7.4%).

A total of 28 interpreting, inferring questions were generated by teachers and pupils. Those 28 questions stimulated 36 responses, 33 of which were at critical levels. A total of 5 evaluating questions were generated by teachers and pupils. Those 5 questions stimulated 6 responses, all 6 of which were at critical levels. A total of 5 gathering specific facts questions were generated. Those 5 questions stimulated 6 responses, 5 of which were at non-critical levels. A total of 2 applying questions were asked. Those 2 questions stimulated 3 responses, all of which were at critical levels. Other questions generated were: 5 clarifying, stimulating one non-critical and one critical response; 3 analyzing, stimulating one non-critical response.

The fact that the number of questions does not correspond exactly with the number of responses is explained by three types of occurrences during the discussions. In two instances the question of the teacher was ignored. This occurred when a student asked a question simultaneously and another student chose to answer the first student's question rather than that of the teacher. In some instances more than one student responded to the same question as students interacted concerning their interpretations of the story. The third type occurence involved students making a response of one type and going on to evaluate or use criteria to support a stand.

DISCUSSION

Results show that, in this study, pupils consistently generated a number of questions equal to or greater than the number generated by teachers. More than half of the questions generated by teachers and pupils were interpreting, inferring questions. Pupils asked more interpreting, inferring questions than any other type. Teachers asked more interpreting, inferring and evaluating questions than any other type.

Pupils' responses to questions indicate that they were thinking critically. More than eighty percent of the responses were critical. The majority of those responses involved interpreting, inferring and evaluating, using criteria and/or supporting a stand.

Findings of this study, like those of Davidson (1978) and Wolf, Huck, and King (1967) indicate that gathering specific fact questions tend to elicit responses which are at a non-critical level. Interpreting, inferring questions tend to elicit

TABLE 1

Summary of Pupils' Non-Critical and Critical Responses by Question Type

		Responses at Non-critical Levels		Responses at Critical Levels		
Asked Questions		Guessing Random	Literal Memory	Inferring Applying Illustrating	Theorizing Hypothesizing	Evaluating Using criteria Supporting stand
Gathering Specific Facts	4	—	5	1	—	—
Clarifying	5	—	1	1	—	—
Interpreting Inferring from facts	28	3	—	13	5	15
Analyzing	3	1	—	—	—	—
Applying	2	—	—	—	1	2
Evaluating	5	—	—	1	2	3

TABLE 2

Summary of Teacher and Pupils' Questions by Question Type: Combined Passages

	Gathering Specific Facts	Clarifying	Interpreting Inferring	Analyzing	Applying	Evaluating	Totals
Teachers	3(6.4%)	2(4.2%)	10(21.3%)	1(2.1%)	1(2.1%)	2(4.2%)	19(40.3%)
Pupils	1(2.1%)	3(6.4%)	18(38.3%)	2(4.2%)	1(2.1%)	3(6.4%)	28(59.5%)
Totals	4(8.5%)	5(10.6%)	28(59.6%)	3(6.3%)	2(4.2%)	5(10.6%)	47(99.8%)

critical level responses. In this study, applying questions and evaluating questions also tended to elicit critical level responses.

This study is considered exploratory. Further investigation is needed to determine whether students and teachers are consistent in the quality, quantity, and variety of questions and responses they generate in discussions following Group Mapping Activities. Further investigation is suggested to determine whether different groups of students demonstrate the same patterns of questions and responses found in this study. Another problem suggested for further study is the question of whether students' and teachers' questioning strategies are influenced by different types of passages.

The Group Mapping Activity is an instructional procedure which provides a means by which teacher and pupil questions are generated and critical responses are stimulated. The results of this study suggest, as Hall has indicated, that creating opportunities for different kinds of thinking processes and true dialogue can lead to students having a "stimulating thinking experience" (Hall, 1979, p. 151).

REFERENCES

Davidson, J. L. The dr-ta: a reading-thinking strategy for all levels. In T. Vacca and J. A. Meagher (Eds.). *Reading as a language experience.* Storrs, Connecticut: The University of Connecticut, 1978, 37-43.

Davidson, J. L. & Bayliss, V. A. Mapping. *Journal of language experience,* 1978, *1*(1), 6-9.

Durkin, D. What classroom observations reveal about reading comprehension instruction. *Reading research quarterly, 14*(4), 481-533.

Guszak, F. J. Teacher questioning and reading. *The reading teacher,* 1967, *21*(3), 227-234.

Hall, M. Ribovich, J. K., & Ramig, C. J. *Reading and the elementary school child* (2nd ed.). New York: D. Van Nostrand, 1979.

Hare, V. C. & Pulliam, C. A. Teacher questioning: a verification and an extension. *Journal of reading behavior,* 1980, 12(1), 69-72.

Smith, N. B. The many faces of comprehension. *The reading teacher,* 1969, *23.*

Stauffer, R. G. *Directing reading maturity as a cognitive process.* New York: Harper & Row, 1969.

Stauffer, R. G. *The language-experience approach to the teaching of reading* (2nd ed.). New York: Harper & Row, 1980.

Taba, H. *Curriculum development—theory and practice.* New York: Harcourt, Brace, & World, 1962.

Wolf, W., Huck, C. S., & King, M. L. *Critical reading ability of elementary school children.* U.S. Department of Health, Education, & Welfare. Office of Education, Project No. 5-1040, Contract No. OE-4-10-187. Columbus: The Ohio State University, 1967.

KAREN K. WIXSON
University of Michigan

THE EFFECTS OF POSTREADING QUESTIONS ON CHILDREN'S COMPREHENSION AND LEARNING

For years educators have advocated the use of different types of questions as an instructional device for producing certain learning outcomes (e.g. Barrett, 1976; Bioom, Englehart, Furst, Hill, and Krathwohl, 1956; Guszak, 1972). Yet, despite the fact that there is a considerable body of literature on the effects of questioning (see Anderson & Biddle, 1975; Andre, 1979; Faw & Waller, 1976), the role different types of questions play in student learning is not well known (Andre, 1979; Kintsch, 1979). This is due in part to the fact that much of the research on questioning is atheoretical (Anderson & Biddle, 1975; Andre, 1979; Kintsch, 1979; Moeser, 1978; Trabasso, 1980). With relatively few exceptions, researchers have utilized informal, intuitive procedures as the basis for identifying question types and for predicting learning outcomes.

Recently researchers have suggested that current schema-theoretic views of the comprehension process hold a great deal of promise for research on the effects of questioning (Andre 1979; Kintsch, 1979; Trabasso, 1980). The essence of such views is that comprehension is a constructive process involving an interaction between the information suggested by the message and the information provided by the comprehender's existing knowledge. Thus, in reading comprehension, meaning is constructed by the reader as a result of an interaction between the "old" information he or she possesses and the "new" information suggested by the text (Adams & Collins, 1977; Anderson, 1977). The integration of old and new information during comprehension results in an adjustment in the reader's existing schematic state. Learning takes place whenever there is a modification in the knowledge base (Rumelhart & Norman, 1978). Therefore, every opportunity the reader has to interact with the textual information is likely to have a direct impact on the reader's comprehension and learning.

Viewed within this context the process of answering a question can be seen as providing the reader with an additional opportunity to interact with the textual information. Accordingly, this additional interaction will have a direct impact on the meaning previously constructed by the reader to account for the text. Further, it is likely that different types of questions will result in different reader-text interactions, resulting in differences in readers' comprehension and learning.

The purpose of this investigation was to determine if different types of postreading questions promote different reader-text interactions, resulting in differences in the nature of children's learning from text. It was hypothesized that different types of postreading questions would promote differences in reader-text interactions which would result in systematic modifications of the knowledge structures previously constructed by the reader to account for the text. Therefore, it was expected that the readers' schematic state at the time of recall would differ from that at the time of comprehension as a function of the type of questions they had received. Specifically, this investigation examined the effects of three different types of postreading questions on the knowledge structures constructed by average and above average fifth grade readers in the comprehension of short, nonnarrative prose passages. In addition, this research was designed to

METHOD

Subjects

The subjects were 111 male and female fifth grade students from a suburban school district near Syracuse, New York. All of the subjects had reading comprehension scores of 5.0 or above on the *Iowa Test of Basic Skills.*

Materials

Seven stimulus passages were selected from *Reading for Concepts-C** (Liddle, 1977; 1970) on the basis of general criteria for length (165-175 words), number of propositions (30-35), readability (4th grade), and on the basis that the passage topics were not included in the school district's science or social studies curricula.

Three types of postreading questions were constructed by the investigator according to the taxonomy of questions developed by Pearson and Johnson (1978). Six open ended questions of each type were constructed for each of the seven stimulus passages. The three types of questions were as follows: (a) textually explicit questions—both the question and the answer are derivable from the text, and the relation between the question and the answer is explicitly represented in the message base of the passage; (b) textually implicit questions—both the question and the answer are derivable from the text, but the answer is implied rather than explicitly represented in the message base of the text; (c) scriptally implicit questions—the questions are derivable from the text, but the answer is neither implied nor explicitly stated in the message base of the text; that is, the reader is expected to provide information from his or her "scripts" or schemata to arrive at the most probable answer. Independent adult raters' agreement with the investigator's classification of 40 questions, randomly selected from the pool of stimulus questions, ranged from .90 to .98 ($M = .93$). The postreading activity for subjects in the control group consisted of six, open ended interest inventory questions constructed by the investigator.

Procedure

Each subject participated in two group testing sessions in the cafeteria within their respective school buildings. Subjects were randomly assigned to stimulus passages and postreading conditions. During the first session each subject read one stimulus passage and wrote the answers to one set of six postreading questions. Subjects assigned to the three experimental groups received either textually explicit (TEQ), textually implicit (TIQ), or scriptally implicit (SIQ) postreading questions. Subjects assigned to the control group received postreading questions which were unrelated to the stimulus passages.

During the second testing session, which took place one week following the first session, the subjects were instructed to, "Write everything you can remember about what you read last week, and anything else it made you think about." In addition, they were asked to write the answers to six, general, open ended questions constructed by the investigator and designed to elicit any additional information they had regarding the passage.

---

*Used by permission of McGraw-Hill, Inc.

Scoring

The stimulus passages and the readers' free and probed recalls were subjected to propositional analysis according to procedures developed by Frederiksen (1975) and modified by Tierney and his colleagues (Bridge, 1977; Tierney & Bos, 1979; Tierney, Bridge & Cera, 1979). Recall propositions were then classified according to a scoring system developed by the author and T. Kosoff for use in this and a related investigation (Kosoff, 1980). The scoring categories were designed to represent different types of information included in the knowledge structures of the readers, and were based on the categories of the Pearson and Johnson (1978) taxonomy used for the construction of the postreading questions. By comparing the text base of the subject's recall to the message base of the stimulus passage, each recall proposition was classified as one of the following: textually explicit (TE); textually implicit (TI): scriptally implicit (SI); or implausible (IMP). Also, the TI and SI measures were added to create an inferred (INF) variable. In order to provide a more comparable index of recall across subjects, the frequencies of each type of recall proposition were converted to proportions of each subject's total recall.

Two raters independently analyzed a total of 20 randomly selected recall protocols. Interrater agreement for the construction of the text bases from written recalls was .97. Agreement for the classification of recall propositions according to the previously described categories was .88.

RESULTS AND DISCUSSION

The data analysis consisted of exploratory analyses (Tukey, 1977; McNeil, 1977), multivariate analysis of variance (MANOVA), with univariate tests (ANOVAs) on each dependent variable followed by planned comparisons and post hoc pairwise comparisons (Dunn's procedure). The primary dependent variables were the relative proportions of textually explicit (TE), textually implicit (TI), scriptally implicit (SI), and implausible (IMP) propositions present in the recall of each reader. An additional variable was the proportion of inferred (TI + SI = INF) recall propositions.

Preliminary data analyses using MANOVA procedures indicated there were no reliable differences in the set of primary recall measures (TE, TI, SI, IMP) as a function of either sex, test group, or passage. Therefore, the data were combined for all further analyses.

The results of a one-way MANOVA with postreading condition as the between subjects factor (TEQ, TIQ, SIQ, CONTROL) and the set of primary recall measures (TE, TI, SI, IMP) as the dependent variables were statistically significant, $F(12,275) = 1.86$, $p < .05$ (Wilk's criterion). The canonical correlation between the condition factor and this set of recall measures was .31, $\chi^2 = 11.01$, $p < .05$. The estimate of the effect size for postreading condition obained from R_c^2 was .10. Thus, as predicted, different types of postreading questions did result in different modifications of the knowledge structures constructed by the children during comprehension.

Planned comparisons were conducted to test for differences in the relative proportions of each type of primary recall proposition between the groups receiving the experimental postreading questions and the control group. The results of these tests revealed that the knowledge structures generated by the subjects who received the experimental postreading questions contained significantly more textually explicit information, $F(1,107)=3.74$, $p<.05$, and significantly less scriptally implicit information $F(1,107)=4.99$, $p<.05$, than the knowledge structures constructed by the subjects in the control group. There were no reliable differences in the proportions of textually implicit or implausible information generated by the subjects in the treatment groups versus those in the control group.

Additional comparisons revealed that the knowledge structures which resulted from the children's interactions with the different types of experimental postreading questions contained similar proportions of textually explicit, textually implicit, scriptally implicit and implausible information. However, the knowledge structures which resulted from the interaction with the explicit postreading questions (TEQ) included significantly less inferred information (INF) than the knowledge structures which resulted from an interaction with either the implicit postreading questions (TIQ, SIQ), $F(1,107)=10.36$, $p<.01$, or the control questions (Dunn's procedure, $\alpha=.05$). There were no reliable differences in the proportions of inferred information generated by the subjects who received the implicit postreading questions and those in the control group. Thus, the subjects in the two implicit question groups and the control group generated similar proportions of inferred information, but significantly more inferred information than the subjects in the explicit question group. These results indicate that the type of postreading question received by young readers influences the degree of interaction between old and new information. Specifically, explicit postreading questions result in significantly less integration of the information suggested by the text with the reader's existing knowledge than either implicit questions or questions which are unrelated to the passage.

An additional analysis was carried out using a two-factor, repeated measures, split plot design with postreading condition (TEQ, TIQ, SIQ, CONTROL) as the between subjects factor in each case and the amount of inferred information, as measured by the levels of explicit (TE) and inferred (INF) information generated by the readers, as the within subjects factor. The results of this analysis revealed that, on the average, subjects produced a significantly greater proportion of inferred information than explicit information, $F(1,107)=349.35$, $p<.0001$. However, the results also revealed a significant interaction between postreading condition and the amount of inferred information produced by the reader $F(3,107)=3.64$, $p<.01$. These results are consistent with previous findings which indicate that the nature of the information contained in the children's knowledge structures at the time of recall differs as a function of the type of postreading activity they received.

The results of this investigation indicate that different types of postreading questions do result in systematic modifications of the knowledge structures constructed by children during comprehension. These results provide further evidence for the view that schemata are assimilative patterns with chronologically recent additions exercising some form of dominance (Bartlett, 1932; Spiro, 1977). When learning is viewed as a modification in the knowledge base (Rumelhart & Norman, 1978), differences in the readers' schemata at the time of recall may be interpreted as differences in learning.

Evidence from this investigation also provides additional support for the view that the degree of interaction between the reader and the text is a continuous variable which is influenced by contextual factors such as task demands (Spiro, 1977). The results indicate that explicit questions result in significantly less interaction than implicit questions. In fact, explicit questions resulted in significantly less interaction than questions which were completely unrelated to the text, suggesting that explicit questions actually inhibit children's integrative processing of textual information. When one considers the current belief that instructional conditions which promote integrative processing facilitate readers' comprehension and learning (Bransford, Nitsch & Franks, 1977; Spiro, 1977; Walker & Meyer, 1980) and the current practice of presenting readers with four explicit questions for every one implicit question (Guszak, 1972), these findings appear to be of major importance.

ACKNOWLEDGMENT

This report is based on a dissertation submitted to Syracuse University in partial fulfillment of the requirements for the Ph.D. degree.

REFERENCES

Adams, M. J. & Collins, A. *A schema-theoretic view of reading* (Technical Report No. 32). Urbana, IL: University of Illinois, Center for the Study of Reading, April 1977. (ERIC Document Reproduction Service No. ED 142 971)

Anderson, R. C. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1977.

Anderson, R. C. & Biddle, W. B. On asking people questions about the way they are reading. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Andre, T. Does answering higher level questions while reading facilitate productive learning? *Review of Educational Research*, 1979, *49*, 280-318.

Barrett, T. C. Taxonomy of reading comprehension. In R. Smith & T. C. Barrett (Eds.), *Teaching reading in the middle grades*. Reading, MA: Addison-Wesley, 1976.

Bartlett, F. C. *Remembering: A study in experimental and social psychology*. London: Cambridge University Press, 1932.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). *Taxonomy of educational objectives: cognitive domain*. New York: David McKay, 1956.

Bransford, J. D., Nitsch, K. E. & Franks, J. J., Schooling and the facilitation of knowing. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1977.

Bridge, C. A. *The text-based inferences generated by children in processing written discourse*. Unpublished doctoral dissertation, University of Arizona, 1977.

Faw, H. W. & Waller, T. G. Mathemagenic behaviors and efficiency in learning from prose materials: review, critique and recommendations. *Review of Educational Research*, 1976, *46*, 691-720.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, *7*, 371-458.

Guszak, F. J. *Diagnostic reading instruction in the elementary school*. New York: Harper & Row, 1972.

Kintsch, W. On modeling comprehension. *Educational Psychologist*, 1979, *14*, 3-14.

Kosoff, T. O. *Effects of written introductions on children's comprehension of text and acquisition of knowledge*. Dissertation in preparation, Syracuse University, 1980.

Liddle, W. (Ed.) *Reading for concepts—C*. St. Louis: McGraw-Hill, Inc., 1977, 1970.

McNeil, D. R. *Interactive data analysis*. New York: John Wiley & Sons, 1977.

Mosser, S. D. Effect of questions on prose unitization. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, *4*, 290-303.

Pearson, P. D. & Johnson, D. D. *Teaching reading comprehension*. New York: Holt, Rinehart & Winston, 1978.

Rumelhart, D. E. & Norman, D. A. Accretion, tuning, and restructuring: three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition*. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1978.

Spiro, R. J. Remembering information from text: the "state of schema" approach. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1977.

Spiro, R. J. Accomodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 84-95.

Tierney, R. J. & Bos, C. S. Personal communication, July 1979.

Tierney, R. J., Bridge, C. & Cera, M. J. The discourse processing operations of children. *Reading Research Quarterly*, 1978-79, *14*, 539-573.

Trabasso, T. *On the making of inferences during reading and their assessment* (Tech. Rep. No. 157). Urbana, IL: University of Illinois, Center for the Study of Reading, January, 1980.

Tukey, J. W. *Exploratory data analysis*. Reading, MA: Addison-Wesley, 1977.

Walker, C. H. & Meyer, B. J. F. Integrating information from text: an evaluation of current theories. *Review of Educational Research*, Fall 1980, *50*, 421-437.

JERRY ZUTELL
Ohio State University

COGNITIVE DEVELOPMENT, METALINGUISTIC ABILITY AND INVENTED SPELLINGS: COMPARISONS AND CORRELATIONS

Recent research (e.g. Read, 1971; Beers, 1974; Beers & Henderson, 1977; Zutell, 1979) has shown that young children's spelling attempts follow an orderly, sequential progression dependent on the sophistication of the child's phonological system, his familiarity with print and his ability to integrate other sources of information (letter patterns, syntax, semantics) into his understanding of the writing system.

Initially children have only a general sense of the relationship between the message expressed by a text and the stream of sounds and letters that convey that message. Attempts at individual words result in what appear to be random strings or, a bit later on, single salient letters (Henderson, 1980). As they are exposed to written language and given the opportunity to manipulate and create their own messages and letter patterns they use the letter-name strategy 'discovered' by Read and since confirmed in several studies (Read, 1971; Beers, 1974; Beers and Henderson, 1977; Zutell, 1979, etc.). At this stage children focus directly on matching individual phonemes with the names of the letters they know. (See especially Read, 1971, for a full explanation of letter-name patterns.) During the next 'transitional' stage children begin to incorporate other aspects of the spelling system into their attempts: correct short vowel spellings stabilize, vowel and past tense markers and silent letters begin to appear, are overgeneralized and then stabilize. In the more advanced (and more protracted) stages children and adults utilize morphological and etymological information in their spellings as well.

This line of research indicates that learning to spell is a conceptually demanding, developmental process which often requires the speller to integrate various aspects of the word (letter-sound name relationships, letter patterns, part of speech, meaning) in producing an accurate spelling. As such, learning to spell could very well involve processing operations similar to those involved in general cognitive growth and in other developmental aspects of literacy learning. For example, Piaget (1973) describes the changes from preoperational to operational thinking partly in terms of the ability to integrate part-whole relationships. (However, see Donaldson, 1978, for some recent criticisms of Piaget's methods and conclusions.) Similarly, Popandropolou and Sinclair (1974) have applied Piaget's method to the study of how young children think about language, particularly their concept of word. Their interviews reveal:

> . . . a slow development of this concept from an initial indifferentiation between word and thing (at the level of metalinguistic thought and not at the level of know how) towards the capacity of conceiving the word both as a meaningful constituent element of larger units and as a unit which is itself built up from smaller elements. (p. 241).

In a recent study Zutell (1979) did find a strong significant correlation (.56) between children's performance on conservation tasks and their spelling strategies, but few other studies have dealt with this issue. Furthermore, the finding of significant relationships among these dimensions of learning have important educational implications. Positive significant correlations between conservation tasks, concept of word and spelling strategies would support the argument that learning to spell is at least partly a conceptual, developmental process. If so, then designers of effective spelling programs should consider the developmental needs, abilities and understandings of the speller in determining when, how and what words and word families are to be taught.

Purpose

The purpose of this study was to explore the relationships between the three dimensions of young children's cognitive development, metalinguistic ability and their spelling strategies. In particular it addressed the following questions:

1) Are there significant correlations between children's scores on conservation, picture integration and metalinguistic tasks and qualitative ratings of their spelling attempts?
2) How well does children's performance on conservation, picture integration and metalinguistic tasks predict their spelling attempts?
3) How do the patterns of these correlations change with grade level (kindergarten vs. first grade) and/or time (beginning vs. end of school year)?

Sample

Sixty-five children were the subjects in this study, 33 kindergarteners and 32 first graders, approximately half of them boys and half girls. All of the subjects attended one of two schools in the Columbus, Ohio metropolitan area, and approximately one-half of the sample at each grade level came from each school. Since both schools used a system of cross-age grouping in their classrooms, students were randomly selected from several different classes within each school.

Both schools were also relatively similar in terms of socioeconomic-status of subjects' families and in instructional philosophies. The children in this study were middle and upper-middle class children attending alternative informal schools. The focus of activities in both schools included many manipulative tasks, reading to children, creative writing, art and drama. Individualized reading selections, rather than basal or programmed materials, were used as the basis of the reading-language arts curriculum. Phonics skills were not taught in a prescribed sequence on a daily schedule, but skill sessions and activities were assigned for individuals and small groups as the teachers felt they were needed.

Procedures

Each child in the study was tested two times during the school year, in October and April, on four sets of instruments: a battery of conservation tasks, the Picture Integration Test, a metalinguistic interview procedure and a set of spelling words.

The conservation battery included conservation of Mass, Continuous Quantity and Volume, and was conducted using traditional Piagetian procedures (Piaget, 1973). Children were classified as nonconservers, confused-transitional or conservers on each task.

The Picture Integration Test (Elkind, in press) consists of a series of seven pictures in which large objects are constructed from a combination of a class of smaller, individually recognizable objects. For example, one picture is an airplane made from several different vegetables (potato, carrots, etc.). The child is asked to describe what he sees. An answer describing only the parts (e.g. "vegetables") receives one point, one describing the large 'whole' object (e.g. "airplane") two points and one that includes both parts and whole (e.g. "an airplane made of vegetables") three points. Thus each child's score fell in a range from 7 to 21 points.

The metalinguistic interview procedure was adapted from Papandropoulou and Sinclair (1974). An examiner asked the child a series of questions regarding what words are and how they work. These included questions about long and short, easy and hard words. Each child was also asked to make up a new word they had never heard or seen before. For each category the children were asked both to provide examples of words and to explain why their examples fit the class of words under discussion. Each child's responses were carefully noted and recorded. Later these were scored using a scale developed for the purposes of the study. Each child thus received a score for the three categories of Length, Ease and Inventiveness.

Spelling ability was measured using a list-test procedure. First, four word categories were selected for investigation: Short Vowel, Long Vowel, Past Tense and Consonant Doubling (see Zutell, 1979). Then two lists of nine words each were constructed using examples from each category. These were administered to small groups of subjects. Children were encouraged to make their best attempts even if they did not know how to spell a word. They were also given as much time as they needed for each word. Spellings were then scored using a qualitative rating scale adapted from Beers (1974), Zutell (1979) and Henderson (1980). Totals were then tabulated for each category (Short Vowel, Long Vowel, Past Tense, Consonant Doubling).

Method of Analysis

The data were subjected to 11 (one for each dependent variable) analyses of variance with repeated measures, with Grade as the between and Time as the

TABLE 1
Multiple Regression Summary Table for Spelling Categories by Grade and Time

	Grade K		Grade 1	
	Time 1	Time 2	Time 1	Time 2
Short Vowel				
Multiple R	.687	.742	.642	.391
R Square	.472	.551	.412	.154
Level of Significance	.05	.01	.05	ns
Long Vowel				
Multiple R	.660	.710	.654	.442
R Square	.435	.504	.428	.195
Level of Significance	.05	.01	.05	ns
Past Tense				
Multiple R	.637	.639	.566	.335
R Square	.406	.408	.321	.112
Level of Significance	ns	.05	ns	ns
Consonant Doubling				
Multiple R	.662	.674	.651	.346
R Square	.438	.454	.424	.120
Level of Significance	.05	.05	.05	ns

TABLE 2

Correlations of Predictors with Spelling Categories by Grade and time

	Grade K				Grade 1			
	Time 1		Time 2		Time 1		Time 2	
	Simple R - R² Change		Simple R - R² Change		Simple R - R² Change		Simple R - R² Change	
Short Vowel								
Mass	.32*	.10	***		.19	.03	***	
Quantity	.17	.00	.24	.06	.15	.00	.25	.07
Volume	.41**	.08	.21	.01	.23	.03	.14	.00
P.I.T.	.00	.00	.32*	.15	.37*	.12	.19	.03
Length	.39*	.06	.54**	.21	.40*	.17	.08	.00
Ease	.47**	.16	.47**	.06	.12	.00	.08	.03
Invent.	.36*	.05	.53**	.06	.28	.06	.03	.02
Long Vowel								
Mass	.35*	.12	***		.04	.00	***	
Quantity	.34*	.03	.25	.06	.10	.01	.27	.08
Volume	.30*	.02	.21	.00	.19	.03	.14	.00
P.I.T.	.06	.00	.26*	.11	.20	.05	.21	.05
Length	.26	.02	.51**	.19	.47**	.23	.06	.01
Ease	.40*	.12	.47**	.07	.26	.06	−.19	.07
Invent.	.40*	.11	.52**	.06	.15	.05	.12	.00

Past Tense								
Mass	.38*	.14	***		.25	.06	***	
Quantity	.41**	.05	.19	.04	.26	.01	.19	.04
Volume	.33*	.02	.19	.01	.28	.03	.21	.02
P.I.T.	.03	.00	.20	.07	.28	.04	.16	.02
Length	.36	.06	.40*	.11	.16	.03	.08	.00
Ease	.41**	.06	.38*	.05	.10	.03	.11	.02
Invent.	.31*	.04	.56**	.13	.38*	.11	.01	.01
Consonant Doubling								
Mass	.35*	.12	***		.21	.06	***	
Quantity	.36*	.04	.20	.04	.18	.00	.29	.08
Volume	.24	.00	.18	.00	.36*	.08	.19	.01
P.I.T.	−.13	.01	.28	.11	.36*	.08	.09	.01
Length	.36*	.07	.44**	.13	.21	.05	.02	.00
Ease	.45**	.10	.41*	.05	.03	.00	−.07	.02
Invent.	.40*	.08	.55**	.11	.40	.15	.06	.00

* p. 05

** p. 01

*** Omitted to avoid problems with colinearity.
Correlation with Quantity .9

repeated measure. The results of the analyses indicated a strong significant effect for Time, but only one significant main effect for Grade (Continuous Quantity) and one for Grade x Time interaction (Long Vowel). Then the data were subjected to Multiple Regression Analyses in which scores on the conservation measures (Mass, Continuous Quantity, Volume), the Picture Integration Test and the metalinguistic scales (Length, Ease, Inventiveness) were used as predictors of the total scores for each of the four spelling categories (Short Vowel, Long Vowel, Past Tense, Consonant Doubling). Separate analyses were done for each grade level and time (Kindergarten, Time 1; Kindergarten, Time 2; First Grade, Time One; First Grade, Time 2). Thus 16 separate analyses were done.

Each analysis yielded a multiple correlation of predictors to spelling score (Multiple R), a measure of the variance in spelling scores accounted for by the predictors (R Square) and a test of statistical significance of the multiple correlation. These data are reported in Table 1. Specific simple correlations between individual predictors and spelling category scores and the *unique* amount of variance accounted for by each predictor are reported in Table 2.

Results

The spelling patterns of the children in this study were similar to those found in other recent studies of children's spelling attempts, supporting a developmental, conceptual model of children's word knowledge. A high percentage of spellings followed letter-name patterns. As expected, Short and Long Vowel categories were considerably easier than Past Tense and Consonant Doubling. For the easier categories in particular, children's strategies were progressively more sophisticated for both grades at Time 2.

As the data in Table 1 indicate, in 10 of the 16 regression analyses the multiple correlation is significant at least at the .05 level, with roughly 40% to 50% of the variance in spelling accounted for by the metalinguistic and cognitive measures. Correlations are particularly strong for kindergarten children at the second collection, and especially for Long and Short Vowel categories. On the other hand, none of the correlations for the first graders at the second collection are significant. In general this relationship holds uniformly for all four spelling categories, with the exception of Past Tense which shows the weakest correlations throughout.

While the summary statistics in Table 1 show a rather consistent pattern of relationships, the individual correlations in Table 2 indicate changing patterns across grade and time. For Kindergarten at Time 1 both conservation and metalinguistic measures are significantly correlated with all of the spelling categories. At Time 2, however, conservation correlations are not significant, while metalinguistic correlations have generally grown stronger. Interestingly, correlations for the Picture Integration Test increase markedly from Time 1 to Time 2. Even though these correlations are extremely low at Time 1 and only significant for one of the four spelling categories (Short Vowel) at Time 2, at the second collection the P.I.T. scores alone are accounting for 7-15% of the variance accounted for in the analysis.

For Grade 1 the patterns of relationships are quite different. At neither collection period do conservation measures significantly correlate with spelling scores. Metalinguistic scores account for most of the variance at Time 1, yet particular correlations (Length to Long Vowel and Short Vowel; Inventiveness to Past Tense and Consonant Doubling) are quite strong and account for much of the variance, while the others are much weaker. P.I.T. scores are significantly correlated with Short Vowel and Consonant Doubling scores at Time 1 but not with Long Vowel and Past Tense. Finally, at Time 2 there are no significant correlations between any of the cognitive or metalinguistic measures and any of the spelling categories.

As predicted, there were significant correlations between cognitive and metalinguistic measures and the spelling variables. For three of the four grade by time samples, a combination of these measures predicted roughly 50% of the variance for each of the spelling categories. In general the metalinguistic measures were the stronger predictors. But ceiling effects for the two easiest cognitive measures (Mass and Quantity) for three of the four samples limited the possibility of significant correlations. The same was true for two of the metalinguistic measures (Length and Ease) for Grade 1 at Time 2. While the more difficult cognitive and metalinguistic measures (Volume, P.I.T., Inventiveness) were on occasion significantly correlated with the more difficult spelling tasks (Past Tense, Consonant Doubling), familiarity with print, the effects of schooling, the ceiling effects on the other variables and the children's generally low level of spelling strategies were evidently more important factors, especially for Grade 1 at Time 2.

Conclusions

Given the above results it seems reasonable to conclude that there is a commonality among general cognitive development, metalinguistic concept of word and the stages of children's spelling ability. Development along all three dimensions involves the ability to 'decenter,' that is, the ability to deal with more than one aspect of a situation or object simultaneously, to view that situation or object in terms of its parts, as whole and as a part of other, larger contexts. The relationship between the three dimensions is strongest at the younger ages just as the children make the transitions from preoperational to operational thinking, from viewing word and object as one to recognizing words as independent objects and from letter-name to more sophisticated spelling strategies.

In terms of instruction, the results of this study suggest that teachers should be cautious about requiring correct spelling and/or sophisticated word analysis skills from young children before they are conceptually ready to handle such tasks. Recent trends to initiate *formal* reading-spelling instruction in pre-school and kindergarten, especially when such instruction is largely skill-based, drill work, seem particularly open to criticism. On the other hand, less formal language based activities which give children the opportunity to manipulate and experiment with oral and written language (e.g. creative writing, inventive spelling) seem more in keeping with a developmental model.

ACKNOWLEDGEMENTS

The author is particularly grateful to Robert Coughlin, Barbara Keifer, Leslie Mass and Mary Beth Webler for the many hours they spent rating children's protocols, and to Marilyn Reed, Pat Caston and the teachers who participated in the study for their generous and enthusiastic cooperation.

The research reported in this paper was supported by a University Small Grant from the Ohio State University.

266

REFERENCES

Beers, J. *First and second grade children's developing orthographic concepts of tense and lax vowels*, Unpublished doctoral dissertation. University of Virginia, 1974.

Beers, J., and H. Henderson. A study of developing orthographic concepts among first graders. *Research in the Teaching of English*, 1977, *11*, 133-148.

Donaldson, M. *Children's Minds* New York: Norton Book Company, 1978.

Elkind, D. *Picture integration tests*. (in press).

Henderson, E. Word knowledge and reading disability. In *Developmental cognitive aspects of learning to spell: a reflection of word knowledge*. Henderson, E. & Beers, J. (eds.) International Reading Association, 1980, 138-48.

Papandropoulou, I., and H. Sinclair. What is a word? experimental study of children's ideas on grammar. *Human Development*, 1974, 17, 241-258.

Piaget, J. *The child and reality: problems of genetic psychology*. New York: Viking Press, 1973.

Read, C. Preschool children's knowledge of English phonology. *Harvard Educational Review*, 1971, *41*, 1-34.

Zutell, J. Children's spelling strategies and their cognitive development. In *Developmental and cognitive aspects of learning to spell: a reflection of word knowledge*. Henderson, E. & Beers, J. (eds.) International Reading Association, 1980, 52-73.

Zutell, J. Some psycholinguistic perspectives on children's spelling. *Language Arts*, 1978, *55*, 844-50.

Zutell, J. Spelling strategies of primary school children and their relationship to Piaget's concept of dencentration. *Research in the Teaching of English*, 1979, *13* (1), 69-80.

WILLIAM H. TEALE
ELETTE ESTRADA
ALONZO B. ANDERSON
University of California, San Diego

HOW PRESCHOOLERS INTERACT WITH WRITTEN COMMUNICATION

Children in the United States bring to school with them their experience of living in a literate society. The general aim of the research discussed here is to characterize the interactions with written communication which a sample of low income preschool children have so that we might understand what these children learn in reading and writing as well as the ways in which they learn it.

The approach adopted in this work is a socio-cultural-psycholinguistic one. We view reading and writing as socio-cultural in nature because we believe that (a) written language is a socially created symbol system which changes with the form of society and the level of its cultural development (Vygotsky, 1962, 1978), (b) that one becomes literate in interactions with others (Teale, 1981), (c) that literacy is a component of social structure (Reder & Green, 1979) and (d) that the nature of literacy is determined by the norms and needs of a particular cultural group (Heath, in press; Reder & Green, 1979; Scollon & Scollon, 1979; Scribner & Cole, 1978; in press). And we conceive of reading and writing as psycholinguistic because we view literacy to be essentially meaningful communication and thus see the need to study it not as an abstraction but rather in terms of the processes involved in reading and writing (Goodman, 1976; Smith, 1978).

These assmptions have direct bearing on our methods of studying how literacy develops in preschoolers. We are interested not only in the 'internal' processes of the individual child as he or she engages in reading- or writing-like situations but also in the ways in which the child interacts with others in his/her daily environment in the social construction of literacy events. We study literacy, then, in the contexts in which it occurs in order to understand how the contexts are organized and how particular practices relate to the child's development of literacy. Naturalistic research techniques are of primary importance for they enable us to pay special attention to the ways in which literacy environments are constructed in everyday interactions between the children and the individuals around them who are more skilled in reading and writing.

SAMPLE

The sample for the first year of the study included twelve low income children and their families, four families representing each of three ethnic groups—Anglo, Black and Mexican-American. Low income children are the focus of the study because, as a group, such children (especially if they are low income children of color) experience particular difficulty learning to read and write in school. We hoped to gain insight into why this is the case by investigating the experiences of living in a literate society which these children bring to school with them. Three ethnic groups are represented in the sample because, in addition to our attention to socio-economic factors, we are also interested in cultural factors which bear on preschool children's socialization into literacy.

During the year we were able to work with only 10 of these 12 children long enough to obtain reportable results. Therefore, the remainder of this article will focus on these children:

	Anglo	Black	Mexican-American
Sex (Age at outset)	M (3.1)	F (3.7)	F (3.5)
	F (3.7)	F (3.6)	M (3.3)
	F (2.4)	F (2.8)	M (2.9)
			M (2.5)

The ten families had from 3 to 7 members (see Table 1 for details), and their incomes were all below $10.000. None of the adults in the families had earned more than a high school diploma.

METHOD

In order to characterize the children's interactions with written communication and what they learn in reading and writing, we employed three data gathering techniques: (1) naturalistic observation, (2) self-report, and (3) controlled behavior samplings.

Naturalistic observations. Each family in the study has been the subject of four hours of observation per week. The observations are intended to be as unobtrusive as possible, the researcher initiating no interaction with members of the family but responding as appropriate. Observations are varied across times of the day and days of the week so that as representative as possible a picture of the daily activities of the target child might be obtained. These observations inform us about the family and community contexts within which literacy events occur.

The focal point of the data collection in the observations is the literacy event, any occasion upon which a person produces, comprehends, or attempts to produce or comprehend written language. Our description of a literacy event includes the activities which lead up to, happen subsequent to and co-occur or alternate with the reading and/or writing itself. And, of course, we detail as completely as possible the reading or writing behaviors. From descriptions of this type we attempt to draw conclusions about the contexts which are literacy events for our families.

Self-report. One caretaker in each family was asked to keep an audio-tape diary of all literacy events that the target child was seen to take part in on the days that the researcher was not observing. These self-reports provide an indication of what occurs in the literacy environment when we cannot be present.

Controlled behavior samplings. The controlled behavior samplings present a common set of literacy experiences to all subjects in the study. The children's behaviors in these experiences provide us with information about individual and cultural diversity in response to situations involving reading and writing. These "interviews" attempt to assess the children's knowledge about writing (Luria, 1929/1977-78; Ferreiro, 1978), book handling (Clay, 1972; Goodman & Cox, 1979), and print awareness (Goodman & Cox, 1979).

RESULTS

Because data from the audio-taped diaries and interviews require a longer collection period before they can be analyzed, the results discussed in this paper stem only from what has been found in the naturalistic observations.

Literacy Materials. One of the first things described from the observations is the literacy materials available to each of the families in the study. These materials are of interest because the types of materials which the adults (and older siblings) use and to which our target children have access set important parameters for the overall literacy environment of each family. Materials in both the home and the community are taken into account.

Table 1 summarizes the materials found in each child's surroundings. It can be seen that there are great differences among families, both in the amount and in the types of materials present. Environmental print—product labels, signs in the community, etc.—is quite similar across children. However, examples of connected discourse vary widely. Some families (Anglo 1, 2, 3; Black 2; Chicano 4) have a variety of books appropriate for and available to the target children. Others have few or none.

The amount and types of adult material vary also. In general, relatively few adult books (either fiction or non-fiction) can be found in the homes. An exception to this finding is Anglo Family 3, where the mother, though she did not own many books, made use of friends and the library to always have a novel available for her personal reading. Other adult materials such as magazines and newspapers appear with greater frequency, though these also are not in great supply. Only the guide to television programs occurs with regularity in the homes. In three of these families the reading materials brought home from school represent virtually all of the textual materials present.

Finally, we have examined the writing materials available to the children. In all of the homes there are some writing instruments and some paper. It is common in most families, however, that when the target children want to write or draw, a pen/pencil or, more often, paper is difficult to locate and must be searched for. The exception to this is Anglo family 2 which has organized a special place for the child's writing instruments and always has plenty of paper available. Thus, ease of access to writing materials varies across families. Furthermore, it appears that ease of access correlates with amount of time spent writing by the target child.

The finding that this correlation exists leads to consideration of the issue of salience of materials. Table 1 describes the literacy materials present in each child's environment; our observations have led us to hypothesize that presence and salience are not necessarily synonymous. The distinction between the two can be illustrated by referring to Anglo Family 1, the Rookers, and Anglo family 3, the Davidsons. The Rookers have in their home, relative to the other families in the study, a large number of literacy materials for the target child. However, these materials do not really play a central role in the environment of the child. They are kept in a box in the bottom of the closet. In comparison the Davidsons have fewer materials present for Kristin, the target child, but the mother has set aside a shelf in Kristin's room for her books. As a result, these materials are more 'in the way'. Interestingly, Kristin's books are a more central facet of her daily activities. Such a situation also exists in Black Family 3 where Kasele, the target child, has her books kept on a prominent shelf in her sister's room. It seems almost that this arrangement 'reminds' Kasele that literacy events are an alternative in which she can participate. Thus, literacy materials for Kasele and Kristin seem to be more salient features of their environments.

In another paper Teale (1981) has argued that a literacy environment is essentially a social rather than a physical phenomenon. The distinction between presence and salience serves as an example of this point. Merely to describe the

Table 1
Literacy Materials in Home Environments.

Anglo 1	*Anglo 2*	*Anglo 3*	*Black 1*	*Black 2*	*Black 3*	*Chicano 1*	*Chicano 2*	*Chicano 3*	*Chicano 4*
Father 22 Mother 22 *Boy 3.10 Girl 1.1	Father 30's Mother 25 *Girl 3.7 Girl 1.4	Father 24 Mother 22 *Girl 2.4	Father 25 Mother 25 *Girl 3.0 Boy 0.3	Father 31 Mother 30 Girl 6.9 *Girl 2.8	Father 38 Mother 32 Boy 15.8 Boy 11.7 Boy 4.8 *Girl 3.7	Father 32 Mother 32 Boy 6.5 Boy 4.7 *Boy 2.5	Mother 45 Father 35 *Boy 3.3	Father 30's Mother 28 Boy 11.3 Girl 8.10 Boy 5.11 Boy 4.10 *Girl 3.5	Mother 26 Girl 5.9 *Boy 2.9
Many children's books (at least 30, esp. Golden Books)	Many children's books (at least 50)	Several children's books (10-20)	Few children's books	Many children's books (20-30 for all children in family)	Few children's books (3 seen)	No childen's books	*Sesame Street Coloring Book*	Only books brought home from school by older children	Many children's books (one box full for each child)
A few paperback (adult) novels	Approx. one dozen novels	Always a paperback or hardback novel around	No adult books	No adult books	Bible, Watchtower pamphlets	No adult books	A few old books (one E.S.L. book)	A few odd books (few novels, mechanics books—some English; some Spanish). Several Spanish novelas.	Several adult books including dictionaries English; English/English Spanish) and books for courses e.g., Real Estate, mother has taken
Few odd magazines	Omni (subscribe) Good Housekeeping (regular basis) Women's magazines	Few odd magazines	Ebony on a regular basis	Ebony (subscription); a few Penthouse	A few old, odd magazines	Two women's magazines (in Spanish)	No magazines	No magazines	Odd copy or two of a magazine
TV Guide	TV Guide	TV Guide	TV Guide	TV Guide	TV listings pamphlet	Use newspaper listings for TV	No TV guides	Home Box Office guide	No TV guides
Newspaper (daily subscription)	No newspapers	No newspapers	Newspaper on a daily basis	No newspaper	No newspapers	Daily newspaper (in Spanish)	No newspapers	No newspapers	No newspapers

Typical labels on products, print on TV etc. Many names on toys.	Typical labels on products, etc. Print on TV often used for literacy learning	Typical labels on products, print on TV, etc.	Typical labels on products, print on TV, etc.	Typical labels on products, print on TV, etc.	Typical labels on products, print on TV, etc.	Typical labels on products, print on TV; majority in English, some in Spanish.	Typical labels on products, print on TV; majority in English, some in Spanish.	Typical labels on products, print on TV; majority in English, some in Spanish.	Typical labels on products, print on TV; majority in English some in Spanish
Typical mail advertisements, few personal letters.	Typical mail plus many personal letters sent and received.	Typical mail plus many personal letters sent and received.	Typical mail, advertisements. Never seen personal letters sent or received.	Typical mail, advertisements. Never seen personal letter sent or received.	Typical mail, advertisements. Never seen personal letter sent or received.	Typical mail, advertisements. Never seen personal letter sent or received.	Typical mail, advertisements. Never seen personal letter sent or received.	Typical mail, advertisements, few personal letters received, sent.	Typical mail, advertisements, occasional personal letters sent.
Many T-shirts with print on them.	Many T-shirts with print on them.	Wallpaper with print on it.	A few T-shirts with print on them.	School books, papers, workbooks, library books. Flyers brought home by older children.	School books brought home by older children.	Flyers, school books brought home by children.	Occasionally neighbor children will bring over basal readers.	Many flyers, books, workbooks home by older older children from school.	Mother occasionally brings home workbook type things from her teacher's aid job.
	Many Literacy-related toy kits, workbooks, etc.			Religious pamphlets	Wallpaper with print on it.	Calendar	5 calendars on walls	Avon catalogues (Mother sells Avon products)	
				Calendar on wall	Calendar on wall				
	Much writing paper crayons, pencils Magic erasable board.								

*Target Child

print which exists around the child is not sufficient for characterizing the child's print environment. It is important to realize that the child actively processes the print around him/her; however it is equally important to realize that print is in varying ways organized and arranged for the child by others. Advertisers attempt to make environmental print in the community salient for all of us. Parents likewise play an important role in affecting the salience of literacy materials in the home. Of course, the creation of saliency of literacy materials is intimately bound with the use of those materials by both the parents and other members of the family and by the target child.

Thus, there exists a great range in the amount and the types of materials in these low-income homes. Were the trends found in this limited sample to be generalized, it would be found that, in comparison with middle class families, low income families, on the whole, encounter typical environmental print to the same degree but have fewer books, newspapers, magazines, writing materials, and so forth available. However, the individual differences which exist among the families must always be kept in mind for they are integral to any attempt to characterize the materials present.

Uses of literacy materials. The description of materials provides only partial insight into the literacy environments of the homes. The *uses* of literacy materials must also be characterized. Analysis of the frequency of literacy events reveals that there is a wide range among families, both in the number of events and in the amount of time spent reading and writing. To enable comparison across families we shall present these quantitative data in terms of average number of events per hour during observation and average time spent in reading and writing per hour.

Among the target children Vicki, the girl in Anglo Family 2 tops the list on all frequency counts. She participated in an average of .89 reading and 1.08 writing events per hour and spent an average of 18.70 minutes of every hour (7.53 minutes of reading and 11.17 minutes of writing) engaged in literacy activities. On the other end of the spectrum is the target child in Chicano Family 2, where the mother is, for all intents and purposes, illiterate. This child participated in an average of .07 reading and .03 writing events per hour for an average time of .67 minutes per hour involved in literacy events.

It could be argued that Vicki's case is a most unusual one, and indeed we feel it is. However, even when the results for Vicki are disregarded, the great range in frequencies still exists among the target children, as is evidenced in this summary:

	Range in average number of events per hour of of observation	Average number of minutes per hour engaged in
Reading	.06 — .48	.15 — 5.66
Writing	.03 — .19	.15 — 2.64

We have also noted the literacy events for which each target child is present but in which she or he does not participate. Gathering this information gives us a more complete indication of the literacy contexts to which the children are exposed in their daily lives. As may be expected from the previous discussion, there is a corresponding range in the amount and types of these activities by the other members of the families. The average amount of time spent in literacy activities ranges from 5.82 minutes per hour (Anglo Family 3) to .53 minutes per hour (Chicano Family 1), with the median being 1.15 minutes per hour.

Literacy activities for most of the target children are engaged in relatively infrequently both on their parts and on the parts of other members of their families. Only in Anglo Families 2 and 3, Black Family 2, and Chicano Family 3 are literacy events relatively frequent occurrences (and the majority of literacy events in Chicano Family 3 are occasions upon which the third grader in the family organizes 'playing school' situations for the other children). Thus, it is generally the case that the quantity of literacy experience in the majority of the low-income homes is not great.

Not only the quantitative aspects of literacy events are necessary for adequate characterization of the nature of the child's literacy environment, however. The assumptions about the nature of literacy upon which this research is based suggest that any attempt to understand the uses of literacy materials must focus also on the ways in which literacy events are organized, viz., on the participants in the events and the ways in which the events are transacted.

A case in point can be found with Vicki, the target child in Anglo Family 2. Vicki is involved in literacy events far more frequently than the rest of the target children, and on the average she spends at least 2.9 times as much time reading and writing as any other child being observed. Yet, from observations and results on the controlled behavior samplings we can see that her literacy conceptions and abilities are not as far advanced as several of our other preschoolers. Why this should be the case is of great interest to us. The difference in the quality of her literacy experiences and her literacy environment from that of other children is being examined, especially through microanalyses of how the target child and other, more 'literacy-experienced' members of the household interact on particular occasions to construct reading and writing activities.

These microanalyses, conducted on representative literacy events from all families give us insight into how effective the more skilled members of the family are at focusing the child toward developmental levels of literacy not yet reached. Of special importance on this latter point is the notion of zone of proximal development (Vygotsky, 1978; Brown & French, 1979). Our hypothesis is that the more effective the parent or sibling is at negotiating the child's zone of proximal development in literacy events, the more effectively the child will learn to read and write.

It is still too early to draw any firm conclusions on these points, however. Literacy interactions that are qualitatively different have certainly been observed in the families. For example, parents read differently to children; the amount of more formal, school-like literacy events varies from home to home. It remains now to develop more completely a means of describing those differences and to assess any effects those differences have on the children.

The functions which literacy serves for these low-income families seem on preliminary analyses to be very 'utilitarian'. When, for example, the adults read or write, it is generally for the purpose of doing something necessary for getting along in everyday life: finding out what programs are on television, making a shopping list, writing a check. There is very little work-related literacy, and for only one parent, the mother in Anglo Family 3, was reading typically an activity for relaxation and pleasure. The same can be said for the non-school related reading and writing of school-age siblings. In the second year of this study the attempt will be to discover what effect this 'utilitarian orientation' has on our target children.

CONCLUSIONS

The conclusions mentioned should be viewed as emerging hypotheses. There is still considerable data collection and analysis which must be performed in this project. However, the following points from the observations will continue to guide the work.

Of importance is the finding that there exists a great range in the home literacy environments of these low-income preschoolers. Materials in the ten homes differ in both type and amount. A few children experience literacy events frequently, the others, comparatively rarely. Thus, to attempt to generalize about *the* literacy background which low-income children bring to school with them would be misguided.

It is clear from the observations that income is not *the* factor which determines a child's literacy experiences. Income may be *a* factor in that families operating on limited budgets may not be able to supply many commercially produced literacy materials for their children; however, there are additional variables which affect the literacy socialization of young children. A child's literacy environment is a complex array of interdependent factors. The written language present in the home and the community is important for it determines the types of discourse with which the child will have practice. The uses of written language are important also. The contexts which give rise to literacy events, the ways in which these events are constructed in interaction, the frequency of events, the literacy levels of parents and older siblings, the motivational factors involved—all form part of the child's literacy set.

The quality of the literacy environments of the ten children in this study does vary. In the one home in this study where the literacy environment would be characterized as a very positive one, we find a mother who, in her daily interactions with her child where written language is involved, is effective at negotiating the child's zone of proximal development. Also in this home literacy skills are practiced in everyday activities. For example, the construction of a shopping list or the baking of cookies would be occasions upon which the child participates in reading and writing rather than simply situations where the adult does all of the literacy work. Moreover, literacy situations in this home are, on the whole, positive experiences for the child. Holdaway (1979) has stressed the centrality of motivational factors to developing a literacy set in preschool children. He contends that when children find rewards in literacy events with parents or siblings, they are impelled into further practice in reading and writing, both with others and by themselves. These positive feelings associated with literacy events are not merely affective, however. What typically happens in our 'literacy-rich' home is that reading and writing involve both emotional and cognitive satisfaction. The child receives praise, attention, affection, and so forth and also participates in what she finds to be interesting intellectual work. Schickedanz (1978) has discussed the relationship between the affective and cognitive in story reading situations, concluding that "a story reading situation that is loaded with positive affect . . . is the same situation that is loaded with information for the child." As she says, "It may be that a child's affective tie to an adult is influenced by the adult's responsiveness to the child's cognitive needs" (p. 54).

Another area which we are investigating is any relationship which may hold between the use of literacy materials and culture. Recent research has demonstrated that the practice of literacy may vary from cultural group to cultural group (Scollon & Scollon, 1979, 1980; Scribner & Cole, 1978; in press; Heath, 1980). The issue we are attempting to shed light on is the degree to which the three cultural groups in this study provide different literacy-related practices for children. Differences in the frequency and organization of literacy events, in the contexts which give rise to such events, and in the functions for which literacy is used will continue to be examined.

As has been mentioned, we are attempting to tease out the factors which seem to promote literacy development in children and to identify the consequences of family and community literacy practices. Through extended observation we have gained insight into the socio-cultural practice of literacy in ten families. By continued work with these and additional families we hope to understand better the everyday literacy environments of low-income children in the United States.

ACKNOWLEDGMENTS

This research is supported by National Institute of Education Grant 79 1035. Colleagues working on this project are Rosalie Bennett, Linda Forrest, Joe Martinez, Shelley Stokes and Billy Vaughn.

REFERENCES

Brown, A. L., & French, L. A. The zone of potential development: Implications for intelligence testing in the year 2000. *Intelligence*, 1979, *3*, 255-273.

Clay, M. M. *Sand.* Auckland, New Zealand: Heinemann Educational, 1972.

Ferreiro, E. What is written in a written sentence? A developmental answer. *Journal of Education*, 1978, *160*, 25-39.

Goodman, K. S. Linguistically sound research in reading. In R. Farr, S. Weintraub, & B. Tone (eds.), *Improving reading research.* Newark, Delaware: International Reading Association, 1976.

Goodman, Y. M., & Cox, V. *The development of literacy in young children.* Proposal submitted to the National Institute of Education, 1979.

Heath, S. B. The functions and uses of literacy. *Journal of Communication*, 1980, *30*, 123-133.

Heath, S. B. Ethnography and education: Toward defining the essentials. In P. Gilmore & A. Glatthorn (eds.), *Ethnography and education: Children in and out of school.* Philadelphia: University of Pennsylvania Press, in press.

Holdaway, D. *The foundations of literacy.* Sydney: Ashton Scholastic, 1979.

Luria, A. R. The development of writing in the child. *Soviet Psychology*, 1977-78, *16*, 65-114. (Translated from: *Problems of Marxist Education*, 1929, *1*, 143-176.)

Reder, S., & Green, K. R. *Literacy as a functional component of social structure in an Alaskan fishing village.* Paper presented at the 78th Annual Meeting of the American Anthropological Association, Cincinnati, Ohio, December 1979.

Schickedanz, J. A. "Please read that story again!" Exploring relationships between story reading and learning to read. *Young Children*, 1978, *33*, 48-55.

Scollon, R., & Scollon, S. B. K. *Literacy as interethnic communication: An Athabaskan case.* Manuscript, Alaska Native Language Center, 1979.

Scollon, R., & Scollon, S. B. K. Literacy as focused interaction. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 1980, *2*, 26-29.

Scribner, S., & Cole, M. Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review*, 1978, *48*, 448-461.

Scribner, S., & Cole, M. *The psychology of literacy.* Cambridge, Mass.: Harvard University Press, in press.

Smith, F. *Understanding reading* (2nd ed.). New York: Holt, Rinehart and Winston, 1978.

Teale, W. *Children learning to read and write: Doing what comes naturally?* Manuscript, Laboratory of comparative Human Cognition, University of California, San Diego, 1981.

Vygotsky, L. S. *Thought and language*, E. Hanfman & G. Vakar (eds. and trans.). Cambridge: The MIT Press, 1962.

Vygotsky, L. S. *Mind in society.* Cambridge, Mass.: Harvard University Press, 1978.

TOBIE S. SANDERS
Ohio State University

THREE FIRST GRADERS' CONCEPT OF WORD AND CONCEPTS ABOUT THE LANGUAGE OF LITERACY INSTRUCTION

Whether or not a child implicitly knows, or overtly can tell a teacher or researcher, what a word is, words are likely to play a central role in the processes that engage children in initial classroom literacy instruction. Traditionally, the language instruction register has been marked by abundant references to words. We hear such directives and comments as: "Touch the first word in the green box on page 12," "That's an easy word, so let's sound it out," and "There are no new words on this page."

A debate has arisen in the literacy learning research that questions whether certain aspects of metalinguistic awareness, including concept of word, and concepts about the language of instruction, are a prerequisite, a facilitator, or a result of successful literacy learning (Downing, 1979; Ehri, 1979; Goodman, 1980; Holdaway, 1979). That debate stimulated the research from which emerged the following examples of matches and mismatches between the language of literacy instruction and three first graders' concepts about that language. In addition to presenting an array of first grader's concepts concerning words and other aspects of language instruction, this study suggests evidence of a complex interactive relationship between aspects of metalinguistic awareness, literacy instruction and becoming literate.

In order to determine if aspects of a child's concepts about the language of instruction were reflected in his/her literacy learning discourse, this study focused upon interactional language that occurred in the literacy learning sessions of a first grade classroom. Ethnographic procedures were utilized to collect language data and describe possible evidence of metalinguistic awareness revealed in the naturally occurring classroom talk. Elicitive procedures were also used in this effort. They included researcher-conducted and near-to-peer conducted interviews of the subjects and informal interviews with the classroom teacher concerning her instructional format and her evaluation of the students' success in becoming literate.

Setting

The classroom site was a self-contained first grade in a midsized school near the inner-city of a large mid-western community. Light industry was interspersed in the school's neighborhood with welfare housing projects and private housing consisting mainly of rental units and mobile homes. The cultural mix of the neighborhood was principally inner-city Black and Appalachian. Court ordered desegregation brought children from an adjoining neighborhood into the school. These children were primarily from the cultural mainstream. The self-contained classroom was typical of many in the school system, although both more and less traditional settings exist in various areas of the city. A basal reading series was used for reading instruction.

Subjects

The teacher in this study was from the cultural mainstream and had six years of experience, all in the site school. She was, therefore, more experienced in working with inner-city Black and Appalachian culture students in her classroom than with those who shared her cultural background.

In addition to the teacher, three children served as subjects in this study. Each of these subjects were male first grade students with the usual kindergarten experience consisting of half-day sessions that focused primarily upon readiness and socialization. One child was from the cultural mainstream, one from inner-city Black culture, and one from Appalachian culture. Cultural membership was determined via family history and performance on sentence repetition tasks. Males were selected because research indicated that in their respective cultures males learn language forms which are more likely to clash with mainstream cultural expections of language than the forms learned and used by females (Abrahams, 1972, 1975; Labov, 1973). Mitchell-Kernan (1969), Trudgill (1975), and Wolfram and [illegible]istian (1976) noted that females' language is more "standard" than that [illegible] males, even in the early years. Both inner-city Black and Appalachian cultures have been described as having more of an oral than literary tradition (DeStefano, Pepinsky and Sanders, in press). Differences in developmental aspects of metalinguistic awareness related to the task of becoming literate were anticipated on the basis of the differences in cultural literacy traditions.

Data Collection

The data in this study were collected by a combination of naturalistic and elicitive procedures. This was consistent with the nature of the research problem, for metalinguistic concepts develop through both student generated and teacher directed experences.

1. *Discourse data*

Discourse of the teacher and the three student subjects during literacy learning sessions in the classroom were collected on audio tapes via wireless microphones worn by each subject. Video tapes of each session were also made. The individual audio tapes facilitated careful transcription and examination of interactional discourse. Video-taped records were available to provide a check against the audio and the opportunity to examine elements of nonverbal interaction as well as elements of verbal interaction.

The discourse data were collected during the third week of school, again at the conclusion of the first grading period and finally at the beginning of the second semester. Although the periods of data collection were comprised primarily of literacy instruction frames, these frames were surrounded by classroom business, such as explanations of seatwork, time structuring, and language used to control individuals and groups of children.

2. *Interviews*

Two interviews, designed to display aspects of metalinguistic awareness, were conducted and recorded with each subject. One set of interviews was conducted by this researcher and one set was conducted by a five-year-old male consultant or informant. The purpose of the near-to-peer interview was to elicit a fuller exposition of metalinguistic insight on the part of the three first graders than might be given to the adult interviewer. An adult's request to be told by a child what one does when one reads appears to be a testing situation as the interviewer asks for information that he or she already knows. However, the questions about how one reads asked by a soon-to-be-first-grader appear more as real questions which are deserving of a full answer.

3. *Literacy profiles*

In addition to the language data and interview, a profile of each subject's success in literacy learning was constructed. The profile reflected scores on two administrations of Clay's (1972) *Concepts About Print Diagnostic Survey, Sand,* written language evaluation, scores on reading series achievement tests, academic progress reports and teacher evaluations of success in learning to be literate.

DATA ANALYSIS PROCEDURES

The classroom literacy learning discourse was transcribed and then analyzed through the Mehan (1979) discourse analysis framework. This is a constitutant analysis framework that is designed to study structuring activities and interaction in the classroom. The framework displayed student initiations and replies as well as teacher initiation, replies and evaluations as distinct interactional components of the discourse. Student-generated components displayed through the Mehan format were then specifically examined for Clark's (1977) six points of evidence of the ability to reflect upon language that may be determined by what children say and do: (1) spontaneous corrections, (2) questions about "correct" words, style and elements of other languages, (3) comments on the speech, reading or writing of others (to include comments on textual format), (4) play with linguistic units, i.e. punning, (5) judgments concerning linguistic structures, and (6) questions about aspects of language.

Teacher-generated components of the discourse were examined for aspects of metalinguistic awareness that the teacher demonstrated as being operant in her literacy instruction.

The subjects' responses to both the near-to-peer and researcher conducted interviews were examined for evidence of metalinguistic awareness. Finally, a comparison was made between evidence of metalinguistic awareness that had been reflected both in the literacy learning discourse data and metalinguistic interviews and the individual subjects' literacy profiles.

The discussion that follows focuses primarily upon the results of the interview procedures; however reference is also made to the results of the analysis of classroom and discourse and literacy profiles of each subject. For a more thorough examination of the discouse analysis and literacy profiles see *Metalinguistic Awareness Revealed in Classroom Literacy Learning Discourse Of Culturally Diverse First Grade Students* (Sanders, 1980).

DISCUSSION

1. *Matches and Mismatches Revealed Through Interviews*

Specific areas of confusion and misconceptions about the language of literacy instruction were revealed in the interview procedures that were not evident in the actual classroom discourse. For example, in classroom discourse the subjects responded correctly to their teacher's initiations of talk about words that begin alike, especially when the teacher provided a choice of responses. Thus, it appeared that their concepts of the language instruction phrase "words that begin alike" matched the instructional purposes of the teacher. The interview procedures, however, revealed that two of the three subject's operant definitions stressed other aspects of likeness. For example, the mainstream child reported that "Dog and God and big and dig begin alike." The Appalachian subject stated that "ball and fall," began alike.

Classroom discourse suggested that the teacher operated upon the notion that a stable concept of word as a structural unit was basic to her literacy instruction program. Researcher conducted interviews, however, suggested that both the mainstream and Appalachian culture subjects had not yet established such a stable concept of word. Rather they demonstrated behavioral evidence of confusion between "a word"and "a letter" as well as confusion between "a long word" and "a string of words." This was indicated in the interview through inaccurate finger pointing to words in well known rhymes. All three of the subjects demonstrated that they were confused by their teacher's use of the phrase "new word" in their reading groups. The subjects indicated that they thought the teacher meant a word that they had never encountered in speech or print. The teacher intended the meaning of "new word" to be one that she had not yet guided them through "sounding out." Interviews also revealed confusion between good work, good behavior, and good language, as well as how you know you are a skillfull reader.

A difference in an awareness of the relationship between using synthetic phonic procedures and reading was evident between what the teacher emphasized in her language of literacy instruction and what the students interpreted her emphasis to mean. Although each of the subjects could respond to specific cues in reading instruction that allowed them to utilize teacher guided synthetic phonics procedures, they had little awareness of how those procedures related to reading. In both near-to-peer and researcher conducted interviews, each of the subjects emphasized that knowing words *is* what reading is all about. Two of the subjects expressed the belief that knowing letter names and their associated sounds was something you have to know in first grade, but that those relationships had little or nothing to do with knowing words. Their reasoning seemed to be constructed as follows: Reading is knowing words. Knowing words means not having to be helped and not having to sound words out. Sound/symbol relationships and other aspects of synthetic phonics are used when you are helped or guided to say a word. Being guided with sound/symbol relationships to say words indicated you do not "know" the word. If you do not know the words, you are not reading.

Neither classroom literacy learning discourse nor researcher conducted metalinguistic interviews revealed evidence of the subject's overt knowledge of pragmatic rules for language use in the social context of literacy instruction. Such awareness however, was abundantly evident in the near-to-peer conducted interviews. Each of the subjects, in their own words, told the young interviewer that in the classroom how you say something is as important as what you say. They also stated that to do really well in first grade you should know the reading words before you get to school. One child actually told the young interviewer, "First grade is like a test. When the teacher asks a question like, "Who knows the first word on this page?", you can just raise your hand and tell her what you already know. Then she will think she taught you." This subject thus reflected a rather poignant sense of audience and getting ahead in the classroom. All three subjects responded to the soon-to-be-first-grader with practical language use information that would, in their opinion, help prepare him for his formal entry into literacy instruction in school.

2. *Differences Between Subjects*

The three culturally diverse subjects provided evidence of differences in the nature of their developing metalinguistic awareness related to concepts about the language of instruction. The differences were not marked by cultural differences as had been anticipated. Rather, the differences appeared to be related to patterns of instruction.

The Appalachian child demonstrated less of a sense of the function that reading and writing serves in the adult world than did the other subjects. However, both this subject and the subject from the cultural mainstream demonstrated similar levels of literacy learning metalinguistic awareness. Both of these students, for example, adhered to an unreliable minimum formula for knowing what a word was. This formula, three letters, plus or minus one letter, was founded by Ferreiro (1980) to be operant in her study of word knowledge conducted with preschool children. Both the Appalachian and mainstream culture students displayed general success on their literacy learning profiles, although there were specific areas of differences in performance.

The inner-city Black subject had the lowest profile of literacy learning success; however, he demonstrated extensive metalinguistic awareness of language related to literacy instruction. He, for example, expressed knowledge that all strings of letters that meet the three letters plus or minus one letter specification are not words and he was able to verbalize the specifics of why or why not certain configurations were words, whether or not he "knew" them. Although this child did not provide any evidence of overt metalinguistic awareness until the second data period, by the conclusion of the study he demonstrated a rather sophisticated set of awarenesses with examples from all six of the Clark (1977) categories.

This subject was in the "low" reading group throughout the academic year. This group did extensive skills work rather than meaningful language work for the first semester of school. No books were used during the first semester. The concept of wordness via sight words was emphasized through the extensive use of flash cards and of dittoed work pages. The dittoed pages displayed little context or meaningful language. Nonetheless, they did draw the subject's focal awareness to metalinguistic aspects of that restricted text as isolated words were marked-out, underlined, and circled by the subject. Thus concept of word as a structural unit was emphasized. This was not the pattern of instruction in other subjects' reading groups.

Instructional procedures may have stimulated the subject to generate isolated bits of metalinguistic awareness related to literacy instruction. These unrelated bits may have been related to this subject's general confusion about the reading process. He knew many of the "how to's" of reading, but he barely knew how to read.

The disparity between the structural concept of word and success in becoming literate for this subject suggested the possibility that teaching in a manner that artifically focuses and develops certain metalinguistic concepts, while ignoring or neglecting the meaningful communication aspects of literacy, may produce a high level of metalinguistic awareness that may indeed inhibit a child from engaging successfully in school literacy learning activities.

The disparity referred to above suggests that not all aspects of metalinguistic awarness advance literacy. Rather, isolated bits of metalinguistic awareness which are learned through instructional programs designed to teach that awareness outside of meaningful texts may result in misfocusing a child's attention during literacy instruction to highly segmented, skills-oriented aspects of literacy learning. Such an over emphasized awareness of the many threads of literacy may hide the patterns inherent in the whole cloth of learning to read and write.

CONCLUSION

In conclusion, this study presents a view of both productive and receptive aspects of classroom language used in literacy instruction from the perspective of three, culturally diverse, first grade boys in a single, though not atypical, classroom. The results of this research suggest an interactive relationship between certain developmental aspects of metalinguistic awareness and becoming literate. Further, this study suggests the hypothesis that metalinguistic awareness generated from a student's involvement with meaningful reading and writing may be less explicit, but far more useful than aspects of metalinguistic awareness generated through activities which remove context from text. Thus the use to which both teachers and learners put aspects of metalinguistic awareness, such as the structural concept of word and other concepts about the language of literacy instruction emerges as a topic for further investigation by both descriptive and more traditional methodologies.

ACKNOWLEDGEMENT

Based, in part, on research conducted under NIE Grant No. G-79-0032 from the National Institute of Education, Johanna DeStefano and Harold Pepinsky, Co-principal Investigators, and by the Mershon Center for Programs on Research and Education in National Security, Leadership and Public Policy.

REFERENCES

Abrahams, R. *A true and exact survey of talking black.* A paper presented at the Conference on the Ethnography of Speaking, University of Texas, April 1972.

Abrahams, R. Negotiating respect, patterns of presentation among Black women. *Journal of American Folklore*, 1975, *88*, 58-80.

Clark, E. V. *Awareness of language: some evidence from what children say and do.* A paper presented at the discussion meeting of the Child's Conception of Language, Niijmeger, The Netherlands, May 1977.

Clay, M. *Concepts about print survey, Sand.* Aukland, New Zealand: Heinnemann Books, 1972.

DeStefano, J., Pepinsky, H., and Sanders, T. Discourse rules for literacy learning in a classroom. In Louise Wilkinson (Ed.) *Communicating in the Classroom.* New York: Academic Press, in press.

Downing, J. *Cognitive clarity and linguistic awareness.* A paper presented at the International Seminar on Linguistic Awareness and Learning to Read, B.C., Canada, University of Victoria.

Ehri, L. Beginning reading from a psycholinguistic perspective: amalgamation of word identities. In

Frank Murry (Ed.), *The Recognition of Words*. Newark, Del.: International Reading Association, 1978.

Ehri, L. Linguistic insight: threshold of reading acquisition. In T. G. Waller and G. E. Mackinnon (Eds.), *Theory and Practice* (Vol. 1). New York: Academic Press, 1979.

Ferreiro, E. *The relationship between oral and written language: the children's view point.* A paper presented at International Reading Association annual convention, St. Louis, Mo. May 1980.

Goodman, Y. *The roots of literacy.* A paper presented at International Reading Association national convention, St. Louis, Mo., 1980.

Holdaway, D. *Foundations of Literacy.* Aukland, New Zealand: Heinnemann Books, 1980.

Sanders, T. *Metalinguistic awareness revealed in classroom literacy learning discourse of culturally diverse first grade students.* Ohio State University, 1980.

JUDITH GREENWALD
Educational Improvement Center
West Orange, New Jersey

CHILDREN'S CONCEPT OF TEMPORAL SEQUENCE AND THEIR COMPREHENSION OF INVERSE SENTENCES

In processing inverse sentences (e.g., "Before I went to the bank, I went to the store," "I went to the bank after I went to the store," "John cried because he fell off the bicycle"), young children initially use an order of mention strategy wherein they assume that the order of mention directly reflects the order of the occurrence of the events described, thus leading to misinterpretation of such statements. Gradually, however, dependence on this strategy disappears and children comprehend the sentences correctly (Amidon, 1976; Amidon & Carey, 1972; Barrie-Blackley, 1973; Beilin, 1975; Bever, 1970; Brown, 1975; Brown & French, 1976; Brown & Murphy, 1975; Clark, 1970, 1973; Clark & Clark, 1968; Ferreiro, 1971; Hatch, 1971; Johnson, 1975; Katz & Brent, 1969; Kuhn & Phelps, 1976; Ritter & Tuinman, 1975; Schmidt & Paris, 1978; Smith & McMahon, 1970).

The basis for this developmental psycholinguistic behavior has yet to be explained satisfactorily. While some researchers refer to chronological development when explaining the gradual disappearance of the order of mention strategy (Amidon & Carey, 1972; Bever, 1970; Clark, 1970, 1973), various researchers designate different ages at which children correctly process inverse sentences independently of the order of mention strategy (Amidon, 1976; Barrie-Blackley, 1973; Bever, 1970; Brown, 1975; Clark, 1970, 1973; Johnson, 1975; Schmidt & Paris, 1978). This lack of consensus in the literature indicates a need to attribute gradual disappearance of the order of mention strategy and ultimate comprehension of inverse sentences to psychological factors in addition to age.

The presence of the order of mention strategy might be explained with reference to the absence of a concept of temporal sequence. If so, comprehension of inverse sentences is achieved with gradual disappearance of the strategy implicating presence of the concept of temporal sequence.

The present study, then, investigated whether comprehension of inverse temporal and causal sentences could be explained by the concept of temporal sequence, as defined by Piaget (1927/1969), independent of verbal aptitude, age, and sex.

METHOD

Subjects

The research participants in this study were five-, six-, and seven-year-olds attending private schools in the New York Metropolitan area. They were 120 native speakers of Standard English who were all progressing normally in school. There were 56 boys and 64 girls in the study all of whom participated with parental permission.

Materials

In this study, verbal aptitude was assessed by the *Peabody Picture*

Vocabulary Test (PPVT), Form B, a measure designed by Dunn (1965). Raw scores were used in all analyses.

A task designed by Piaget (1927/1969) and replicated by Lovell and Slater (1960) was used to measure the child's concept of temporal sequence with particular reference to the child's ability to reconstruct past occurring events. On this task, children received scores of 1 to 5 reflecting stages and substages of the concept as described by Piaget (1927/1969). According to Piaget, children at a preoperational stage are able to follow the flow of time as it occurs but unable to reconstruct it. Children at a transitional stage are able to reconstruct temporal sequence, but only by relying on a notion of spatial sequence while children at an operational stage are able to reconstruct past events independently of notions of spatial succession and solely on a concept of time sequence.

To assess comprehension of inverse temporal and inverse causal sentences, children read five inverse *before* sentences, five inverse *after* sentences, and five inverse causal sentences linked by the connective *because*. Scores for comprehension depended upon the number of sentences responded to correctly by the child. A correct response was one in which the child arranged two pictured events described by the sentences according to the order of occurrence of the events—this despite the fact that the inverse sentences mentioned the events in the reverse order.

Procedure

A screening measure, given three weeks prior to experimental testing, was used to ensure that all children understood forward temporal and forward causal sentences in which order of mention is chronological (e.g. "I went to the store before I went to the bank," "After I went to the store, I went to the bank," "Because John fell off the bicycle, he cried"). Five forward *before* sentences, five forward *after* sentences, and five forward causal sentences with the connective *because* were administered so that the probability of getting all five sentences of each type correct by chance was about 3% or less. Sentences used in both screening and experimental testing were presented in counterbalanced order. Non-readers and children unable to comprehend forward sentences were eliminated from the study.

The 120 children who participated in the study met with the experimenter in individualized settings during regular school hours. Each child participated in two testing sessions—one in which the measure for comprehension of inverse sentences was administered, followed three weeks later by a second in which the *Peabody Picture Vocabulary Test* (Dunn, 1965) and the test of the concept of temporal sequence were administered. Thirty of the subjects were administered the concept of temporal sequence test once more, three weeks later, to determine the test-retest reliability of the task.

RESULTS

Factor analysis revealed that the three sentence categories were not distinct and that they represent a single dimension of inverse sentences. The internal consistency reliability coefficient (Cronbach's alpha) for the test of the comprehension of inverse sentences was .78, confirming the reliability of the test. The internal consistency reliability coefficient computed for the *Peabody Picture*

Vocabulary Test (PPVT) was .93, indicating a high degree of homegeneity of this test population. Finally, the test-retest reliability coefficient for the test of the concept of temporal sequence was .94, indicating a high degree of stability over a three week interval for the Piagetian task.

A hierarchical regression model was employed to analyze the relationship among verbal aptitude, age, sex, concept of temporal sequence, and comprehension of inverse sentences. Table 1 indicates that at the first stage of the regression analysis, about 25% of the variance in comprehension of inverse sentences was explained by verbal aptitude, age, and sex. The multiple correlation (.52) of these three independent variables with comprehension of inverse sentences was significant, F (3, 116) = 13.45, $p < .001$. Further, when concept of temporal sequence was entered into the regression equation, over 50% of the variance in comprehension of inverse sentences was explained by the four independent variables (verbal aptitude, age, sex, and concept of temporal sequence). The multiple correlation among the variables (.73) was significant, F (4, 115) = 32.33, $p < .001$, as well. Most importantly, the concept of temporal sequence accounted for 27% of the variance in comprehension of inverse sentences beyond that proportion of the variance accounted for by verbal aptitude, age, and sex. This increment was significant, F (1, 115) = 43.31, $p < .001$.

Due to the significant, though moderate, correlation of age with comprehension of inverse sentences, (.49), $p < .001$, and the significant, though moderate, correlation of age with concept of temporal sequence, (.55), $p < .001$, a regression analysis was performed to determine whether age accounted for a significant increment in the variance in comprehension independent of concept of temporal sequence.

At the first stage of the analysis, 51% of the variance in comprehension of inverse sentences was explained by concept of temporal sequence. The correlation, (.71), of concept of temporal sequence with comprehension of inverse sentences was significant, F (1, 118) = 127.5, $p < .001$. When the independent variable, age, was entered into the regression equation, 52% of the

Table 1
Results of Multiple Regression Analysis of Verbal Aptitude, Age, Sex, and Concept of Temporal Sequence Predicting Comprehension of Inverse Sentences

Source	**Multiple R**	**R^2**	**df**	**F**
Verbal Aptitude + Age + Sex	.518	.258	3,116	13.45***
Verbal Aptitude + Age + Sex + Concept of Temporal Sequence	.727	.529	4,115	32.33***
Increment due to Concept of Temporal Sequence		.271	1,115	43.31***

***p<.001

variance in comprehension of inverse sentences was explained by the combination of concept of temporal sequence and age. The multiple correlation (.72) among these two independent variables and the dependent variable was significant, F (2, 117) = 65, $p < .001$. However, it is important to note that age accounted for 1.2% of the variance in comprehension of inverse sentences beyond concept of temporal sequence, an increment which was not significant, F (1, 117) = 2.5, $p < .001$.

DISCUSSION

The finding that concept of temporal sequence accounted for a significant increment in the variance in comprehension of inverse sentences independent of the effects of verbal aptitude, age, and sex contributes to the body of research in this area in a number of ways. First, the finding is consistent with literature which has demonstrated that variance in processing other complex linguistic structures can be explained, in part, with reference to the levels of cognitive ability (Anyon, 1976; Beilin & Kagan, 1969; Ferreiro, 1971; Greenfield, Nelson, & Saltzman, 1972; Scholnick & Adams, 1973; Sinclair & Ferreiro, 1970). In addition, the finding suggests that concept of temporal sequence is a statistically significant correlate of linguistic structures in which order of mention does not directly reflect temporal sequence of events described (i.e., of inverse temporal and causal sentences). Weil (1970) was unable to explain comprehension of inverse temporal sentences with reference to a more specific and analogous cognitive ability than a general sense of time. While Ferreiro (1971) explained comprehension of inverse temporal sentences with reference to general cognitive ability (the ability to conserve) and with reference to concept of the succession of perceptible events, she did not submit her findings to rigorous statistical analysis. Thus, the present findings support other research which attempts to explain psycholinguistic development with reference to cognitive development and enriches research which explores specific cognitive correlates of comprehension of inverse sentences.

One cannot definitely conclude from results here that the relationship between thought and language is causal or that training in particular cognitive skills can influence comprehension of particular linguistic structures. Nevertheless, the results do suggest that we need to allow for development of particular concepts before exercising children in verbal propositions which incorporate these concepts in our language. Once children obtain a concept of temporal sequence, for example, they may be encouraged to capitalize on this schema to process inverse sentences. Indeed, the finding of this and other, similar studies that there is a correlation between cognitive ability and language comprehension lends support to recent reading research which suggests that instruction must focus on helping readers access prior knowledge and cognitive skill when comprehending text (Adams & Collins, 1979; Anderson, Spiro, & Anderson, 1978; Pearson, Hansen, & Gordon, 1979).

Finally, it was not the case in this study that inability to perform perfectly on the cognitive task precluded children from comprehending all of the inverse sentences. Recently, researchers have been speaking more about the interaction of cognitive development and linguistic growth than about one type of development unidirectionally influencing the other type (Beilin, 1975, 1976; Moore & Harris, 1978). Thus, the present study supports the theoretical notion, recently articulated by Piaget (Beilin, 1976), that there perhaps is an interaction between cognition and language and that development in both domains may be due to a third, abstract mechanism independent of both cognitive and linguistic skill.

REFERENCES

Adams, M. J., & Collins, A. A schema theoretic view of reading. In R. O. Freedle (Ed.), *New directions in discourse processing* (Vol. II). Norwood, New Jersey: Ablex, 1979.

Amidon, A. Children's understanding of sentences with contingent relations: Why are temporal and conditional connectives so difficult? *Journal of Experimental Child Psychology*, 1976, *22*, 423-437.

Amidon, A., & Carey, P. Why five-year-olds cannot understand *before* and *after*. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 417-423.

Anderson, R. C., Spiro, R. J., & Anderson, M. C. Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal*, 1978, *15*, 433-450.

Anyon, J. *Comprehension and production of passive sentences: Structural parallels in cognitive and linguistic development*. Unpublished doctoral dissertation, New York University, 1976.

Barrie-Blackley, S. Six-year-old children's understanding of sentences joined with time adverbs. *Journal of Psycholinguistic Research*, 1973, *2*, 153-166.

Beilin, H. *Studies in the cognitive basis of language development*. New York: Academic Press, 1975.

Beilin, H. Constructing cognitive operations linguistically. In H. Reese & Lipsitt (Eds.), *Advances in child development and behavior*. New York: Academic Press, 1976.

Beilin, H., & Kagan, J. Pluralization rules and the conceptualization of number. *Developmental Psychology*, 1969, *1*, 697-706.

Bever, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: John Wiley & Sons, 1970.

Brown, A. L. Recognition, reconstruction, and recall of narrative sequences by preoperational children. *Child Development*, 1975, *46*, 156-166.

Clark, E. V. How young children describe events in time. In G. B. Flores d'Arcais & W. J. M. Levelts (Eds.), *Advances in psycholinguistics*. Amsterdam: North Holland, 1970.

Clark, E. V. How children describe time and order. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development*. New York: Holt, Rinehart & Winston, 1973.

Clark, H. H., & Clark, E. V. Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology*, 1965, *33*, 300-308.

Dunn, L. M. *Expanded manual for the Peabody Picture Vocabulary Test*. Circle Pines, Minnesota: American Guidance Service, Inc., 1965.

Ferreiro, E. *Les relations temporelles dans le langage de l'enfant*. Geneve: Libraire Droz, 1971.

Greenfield, P. M., Nelson, K., & Saltzman, E. The development of rule-bound strategies for manipulating seriated cups: A parallel between action and grammar. *Cognitive Psychology*, 1972, *3*, 291-310.

Hatch, E. The young child's comprehension of time connectives. *Child Development*, 1971, *42*, 2111-2113.

Johnson, H. L. The meaning of *before* and *after* for preschool children. *Journal of Experimental Child Psychology*, 1975, *19*, 88-99.

Katz, E., & Brent, S. Understanding connectives. *Journal of Verbal Learning and Verbal Behavior*, 1969, *7*, 501-509.

Kuhn, D., & Phelps, H. The development of children's comprehension of causal direction. *Child Development*, 1976, *47*, 248-251.

Lovell, K., & Slater, A. The growth of the concept of time: A comparative study. *Journal of Child Psychology and Psychiatry*, 1960, *1*, 179-190.

Moore, T. E., & Harris, A. E. Language and thought in Piagetian theory. In L. S.Siegel & C. J. Brainerd (Eds.), *Alternatives to Piaget's critical essays on the theory*. New York: Academic Press, 1978.

Pearson, P. D., Hansen, J., & Gordon, C. The effect of background knowledge on young children's comprehension of explicit and implicit information. *Journal of Reading Behavior*, 1979, *11*, 201-209.

Piaget, J. *The child's conception of time* (A. J. Pomerans, trans.). New York: Balantine Books, 1969 (originally published, 1927).

Ritter, J. H., & Tuinman, J. Understanding of temporal construction by kindergarten children. *Journal of Psychology*. 1975, *19*, 163-171.

Schmidt, C. R., & Paris, S. G. Operativity and reversibility in children's understanding of causal sequences. *Child Development*, 1978, *49*, 1219-1222.

Scholnick, C. R., & Adams, M. J. Relationships between language and cognitive skills: Passive voice comprehension, backward repetition and matric permutation. *Child Development*, 1973, *44*, 741-746.

Sinclair, H., & Ferreiro, E. Etude génétique de la compréhension, production et répétition des phrases au mode passif. *Archives de Psychologie*, 1970, 40 (160), 1-42.

Smith, K., & McMahon, L. Understanding order information in sentences: So[illegible]ent work at Bell Laboratories. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances [illegible]olinguistics*. Amsterdam: North Holland, 1970.

Weil, J. *The relationship between time conceptualization and time language in young children*. Unpublished doctoral dissertation, City University of New York, 1970.

ROSEMARIE PARK
University of Minnesota

A CRITICAL REVIEW OF DEVELOPMENTS IN ADULT LITERACY

The field of adult literacy is large in scope, ill defined and draws from a variety of disciplines. Although literacy programs are not new in the United States, adult literacy as a focus for disciplined research and inquiry is new. Illiteracy was identified as a social problem along with many other social problems in the Sixties. Government recognition of this problem resulted in the Adult Education Act in 1966. Educators and researchers have responded to the Federal interest in adult literacy partly because funds are available, partly because funds in the Elementary and Secondary Education areas are shrinking and partly because the adult literacy area is a new and untried area of research.

The purpose of this paper is to review the research and instruction in adult literacy since the inception of the Adult Education Act of 1966. It will attempt to provide an organizational framework for research in the area, link selected research to current practice in instruction, describe the limitations of the research, and point to areas of future need. Since the area is so ill defined, section one of this paper will address definitions, section two deals with research, and section three is a summary of research with implications for the future. The paper will not address literacy studies conducted outside of the United States or deal with research in the area of English as a Second Language.

PART I: DEFINING THE FIELD

The extent of adult literacy in the United States is partially dependent on how literacy is defined. If the standards for literacy are set high, then a greater proportion of the population will fall short of them. Expectations for literacy in the United States have taken a steady upward trend (Resnick and Resnick, 1977). In colonial times it was the mark of a literate man that he could sign his own name. Today the army is concerned that recruits are ill prepared to read and understand technical manuals and materials accompanying weapons systems. These may total up to a quarter of a million pages of advanced technological reading (Wisher, 1980).

Literacy can be defined either in absolute or relative terms. Examples of absolute measures of literacy include years of schooling (UNESCO's definition of literacy was four years of school), level of achievement on school related standardized tests, successful demonstration of a defined set of competencies such as reading a short paragraph and answering questions. In describing the demographics of adult literacy in its 1974 report on the extent of adult illiteracy, the National Association of Continuing Adult Education (NACAE) was reduced to describing literacy in terms of years spent in school. According to their figures, 53 million adults have less than twelve years of schooling. Twenty-three million have less than five years of schooling. Clearinghouse Offender Literacy (1973) used performance on standardized reading achievement tests which essentially compared performance to the average within any particular grade. They reported an average reading grade level of between fourth and fifth grade for incarcerated

290

youth and adults. The Harris Poll, however, used the ability to read and answer questions about classified newspaper advertisements, telephone book entries and application forms. Fifteen percent of American adults were designated as illiterate by Harris (1971). A similar study by the National Educational Assessment Program (NEAP, 1978) labels 13 percent of high school seniors as illiterate based on their ability to read and answer questions on narrative text.

Relative definitions of literacy try to account for the literacy needs of the individual and the changing needs of a technological society. This "literacy" is defined as the ability of an individual to fulfill those functions in life that he or she has to perform. A college professor may need a high level of content knowledge and reading proficiency in statistics in his or her area of specialization, whereas a short order chef needs minimal reading, writing and computation skills to be functionally literate.

The trend, both in relative and absolute definitions has been toward functional definitions. Northcutt in the Texas Adult Performance Level Study (1975) defines literacy in terms of performance in five "life skill" areas: consumer economics, government and the law, occupational knowledge, community resources, and health. Criterion levels of inadequate, adequate, and superior are based on a socio/economic definition of literacy. To be inadequate in any one area depends on scoring below the mean sample score for those without adequate disposable income, level of education and occupational status (Northcutt et al. 1975). Northcutt identifies 20% of American adults as having inadequate reading skills, 32% as having inadequate computation skills.

In terms of research, absolute definitions of literacy are easier to manage and more available than relative ones. For example, data on number of years schooling received is readily available through the 1970 (soon 1980) Census.

Relative definitions of literacy, although theoretically more satisfying, are virtually unmanageable for the researcher. A literacy level for each task in life would have to be established, and then re-established as tasks changed over time. Establishing the exact level of competency required in each task requires a lengthy validation technique enlisting "experts in the field" who very often differ in their views (Griffith, Cervero, 1977). The Adult Performance Level Project (Northcutt et al. 1975) and other work by Fillbrandt and Merz (1977) have attempted to establish competency levels for adult basic skills. The major practical and theoretical problem with these efforts is establishing an appropriate level for each competency. Is it enough to perform at the same reading level as the average worker doing the task you establish competency for? Is there some level that would ensure maximum efficiency for a given task? Do you train adults to have exact competencies in any one area or do you establish a set of minimum competencies that will have the maximum generalizability to many life skill areas. The effectiveness of the various curricula is discussed in a later section of this paper.

PART II. RESEARCH

Characteristics of the Target Population

There is general agreement by researchers about who the target population for literacy is. Those who are illiterate are generally poor. Minorities, the elderly, the incarcerated are overrepresented in this group (NACAE, 1974). The research catalogues the multiple disadvantages of the poor—lack of education, socio-psychological deficits, and ill health (NACAE, 1974; Anderson and Niemi, 1969). Anderson and Niemi provide the best review of the literature in this area, although it is dated. Jencks et al., in *Inequality* (1972) provide a more sophisticated analysis of the problem. As Jencks et al. point out, the poor suffer in both quantity and quality of schooling. Parenthetically, increasing the years spent in school will not guarantee a better job. Discrimination and family status are also powerful factors in determining who climbs out of the ranks of the poor (Jencks, 1972).

Researchers seem to agree that there is a culture of poverty similar to that suggested by Oscar Lewis in *Children of Sanchez* (1961). Such a culture is characterized by a unique value system which deemphasizes such middle class values as the ability to delay gratification. Loyalty to immediate friends and family rather than society as a whole roots the poor to a lower rating on Kohlberg's scales of moral growth (Kohlberg, 1969). Authors suggest that rigidity of thought patterns (Ulmer, 1969) or belief in luck (Mezirow, Darkenwold, Knox, 1975), predominate among the poor and illiterate.

This research leads to what can only be described as the catastrophic approach to adult literacy. It is overly simplistic and in many instances it does not reflect the data. The average person in a literacy program is a white woman in her twenties, in good mental and physical health. She may have economic worries (NACAE, 1974), but she is not crippled by them. She needs constructive help, not what the literature invokes: pity, paternalism, and guilt.

RECRUITMENT AND RETENTION

Much of the money for research in Adult Basic Education (ABE) comes from the Adult Education Act 309 Funding designated for short term demonstration projects or program research. This has resulted in a wealth of reports on local recruitment and dissemination efforts. Best documented is the Appalachian Adult Education project (Eyster, et al., 1974 on). Eyster identifies four types of adult learner. "the stationary poor" being the least accessible to service by ABE programs. These reports are excellent resource materials for new programs.

Generally, the research seems to indicate that word of mouth and community agencies are most effective in recruiting adults (Rossman, 1971). Teacher race and type of curriculum may also be important factors in retaining students (Darkenwold, 1975). Adults drop from programs mainly because of new jobs (Osso, 1975) but many cite child care difficulties, transportation and illness as well as lack of interest for leaving program (Seaman, 1971). The average dropout rate for adults in Basic Education programs is 33% (Hunter and Harmon, 1979).

Generally, the research provides good descriptive statistics on how people hear of programs, how long they stay and why they leave. It also provides a number of models for program outreach. It still does not indicate, however, why 98% of the target population does not participate in Adult Education. Hunter and Harmon recommend that community agencies other than strictly educational ones be involved in Adult Basic Education. They suggest programs focus more on the most disadvantaged and those who enroll least in programs. They urge an all out effort at dissemination through a variety of agencies. They also point to the need for a more systematic and accurate data base as a vital component in ongoing program evaluation (Hunter and Harmon, 1979, p. 134-135).

Overall, the literature seems to reflect the feeling that illiteracy is a social as well as an education problem. Radical analyses point out that illiteracy, like poverty, will be with us until society changes its values and existing structures so as to eliminate poverty. Perhaps the most recent proponent of this point of view is Johnathon Kozol (1978). He suggests that the United States follow Cuba's example and make the elimination of illiteracy a goal of national priority. Illiteracy can be eradicated if literacy is taught in a socially relevant context. This analysis is a slightly newer version of Warren Zeigler's ideas on civic literacy (Zeigler, 1976).

The role of program content and its importance in keeping, interesting and informing students has received less emphasis than the overall demographic descriptions discussed so far. The conventional wisdom, as reflected in writing and research, now seems to be that the adults *want to* learn functional literacy skills (Flaherty, 1977) and *learn better* using such material (Manning, 1978), especially *if it interests them* (Hutchinson, 1978). Much impetus has been given to this research since the Adult Performance Level Project (Northcutt, 1975).

Actual research in the effectiveness of functional life skills curriculum compared to the more traditional grade school type of academic curriculum in teaching reading and writing is equivocal. A large, poorly controlled study by Dauzat (1977) suggests that the Adult Performance Level functional content is more effective in teaching reading skills and retaining students. Work by Bauer (1980) found no clear cut advantages to either method. The Bauer study found that traditional academic content may be more effective in producing reading gains as measured by both the Adult Basic Learning Exam and the APL tests.

INSTRUCTIONAL METHOD

Increasingly, functional life skills content has been linked with competency based instructional delivery systems. Students are pretested to identify entry level of functional skills. Exit competency is then identified and the student works toward meeting and demonstrating the exist competencies within his or her chosen area, usually at his or her own rate (Spady, 1979). This flexible delivery system has been increasingly used with the Adult Performance Level curriculum and is heavily functional in content.

One of the effective delivery systems for a competency based program is computer aided instruction (CAI). This system has the capability of placing students in a given curriculum at the appropriate level, teaching a given skill at the learner's own pace and testing exit competencies. The system can switch learners to remedial track to reinforce skills and can collect data for teachers on student progress. Research on adult learners does suggest that CAI used with traditional academic content—that is content not tied to functional or life skills—is a very effective means of teaching adults (Wisher, 1980; Caldwell, 1978, Park, 1978). However, more research on a variety of populations beyond Navy personnel and correctional inmates needs to be done.

TEACHER TRAINING

The majority of teachers in the literacy field are untrained in adult education. Most states require teachers to hold an elementary license to teach basic literacy skills. Some states such as North Carolina require no licensure at all. Some states such as New Jersey are moving towards a competency based adult teaching license (Park, 1977). Research by Robinson (1974) suggests that 70% of adult teachers have had no coursework in reading instruction, 65% have not taught reading at any level. Demographic data collected in Nebraska does show, however, that teachers are generally receptive to the idea of continuing education for themselves (Even, 1975). This may be difficult to implement since nearly all the descriptive data of ABE teachers shows them to be part-time employees, paid by the hour (Park, 1977; Spear et al., 1972).

The most extensive teacher training study was done by Spear et al., at the University of Missouri (Spear, 1972). Extensive data were collected on teacher, administrator, and student attitudes. Some interesting differences were noted in what teachers perceived student's needs to be and what students themselves identified as needs. Students tended to rate subject matter knowledge high, whereas teachers were more concerned with empathetic relations with the student.

Mocker et al. (1974), have developed an inventory of teacher competencies for teachers in traditional and APL programs. It allows teachers to identify which competencies in the areas of knowledge, behavior and skill they have and those they feel they most need. With 291 different competencies to rank discriminating between each competency statistically is problematical. The inventory also relies on the teacher's own perception of need rather than on any objective scale.

THE AGED

Recently there has been a resurgence of interest in the aged population, those 65 years of age and older. Program statistics show older adults to be over-represented among the illiterate (Buswell, 1971) and least likely to participate in educational programs. Since one of the major rationales of the Adult Education Act and Literacy training is to teach adults to read so that they can get off welfare rolls and become active members of the work force, the aged have been generally bypassed as being nonproductive. The increased interest in aging as a phenomenon and the increased militancy on the part of senior groups, however, has caused more attention to be given to the aged. At this point there is no evidence of increased participation in literacy programs. This may be a reflection of the lag in the availability of program figures by about three or four years. (See figures given in Hunter and Harmon, 1979, as an illustration of this).

Research on the reading habits of older people suggests that they continue with the reading patterns they developed earlier in life (Sharon et al., 1979). Hasse, Robinson and Beach (1979) suggest that older learners are more cautious and concerned about accuracy. The same report also found older learners more easly distracted (Haase, Robinson and Beach, 1979). Glynn and Muth (1979) found older adults less adept at encoding and retrieving information in higher order conceptual chunks and less able at developing mediational strategies for learning than young adults. Clearly there is a great need for more to be done in assessing the needs of this particular group.

ADULT VOLUNTARY READING

Once adults are literate what do they read? The data reviewed are equivocal. Two Gallup Polls done seven years apart suggest the while 58% of the sample confessed to "never having read or finished a bock" back in 1969, by 1976, 84% had read "all or part of a book within the last year." Sharon (1972) did the most extensive study of adult reading habits in 1972. He found that literate adults spent about 1 hour 46 minutes a day reading. However, there was great variability amongst individuals. Murphy (1975) also reports an average time of 1 hour 46 minutes spent reading each day. Blacks in a study by Kling and Tivenan (1978) read about 52 minutes each day. However, Mikulecky in a study on adult reading habits found race not to be a predictor of reading habits (Mikulecky, 1979). This contradicts Sharon (1973) who did find race to be an important factor.

Newspapers and magazines rate high in reading frequency (Sharon, 1973; Kling and Tivenan, 1978; Murphy, 1975). However, low socio-economic groups give less credibility to these sources of information than to television, according to an earlier study by Greenberg and Dervin (1968).

The major surprise in this area is that education appears not to be a strong factor in predicting the amount of reading people do (Mikulecky, 1979). This finding may be a result of the homogeneous population of Mikulecky's study. The average income of the Indiana groups was between 10 and 20 thousand dollars. Had the socio-economic range been wider, then differences might have been found.

Mikulecky did find that sex was a predictor of reading habits. Men did more work related reading than women. This reflects the greater male participation in the work market. Men also work at higher level jobs than women who work in lower status occupations. This explanation might also account for the differences Kling and Tivenan found between the amounts whites and blacks read. The Kling study reflected a lower socio-economic group who might be expected to do less work related reading.

The methodology of this type of study presents some problems. Asking adults if they have read a book recently produces differing results. This could be function of the way the questions have been worded. Most recently, public opinion polls on the abortion issue have come up with very conflicting results that have been shown to be a function of the different ways the questions have been worded. All three attempts at opinion polls on reading have come up with varying results (Gallup, 1969, 1976; Yankelovich, 1978).

One additional problem is the accuracy with which adults view their own reading habits. Much reading on buses, at work, around the house is semiautomatic in nature and passes almost unnoticed. Conscious reading of newspapers or books may vary considerable throughout the year. Book reading may increase during vacation time or on beaches. Similar problems of adults' awareness of the reading they do has been faced by those who attempt to document reading done in vocational settings to be reviewed in the following section.

Much of the research in vocational literacy can be attributed to the interest of the military in raising reading levels within its ranks. In 1968, Project 100,000 resulted in a downward revision of the Armed Services qualifying standards. This led to an influx of marginally literate personnel particularly in the Army (Sticht, 1975). A major three-pronged approach was initiated by the Army to develop methodologies for determining functional literacy levels of various military jobs, to explore techniques for reducing these literacy demands, and to develop programs to teach functional literacy skills appropriate to the minimum demands of the job (Sticht, 1975).

Most of this research has been done by Tom Sticht and his colleagues at Human Resources Research Organization (HumRRO). Essentially, Sticht et al.

TABLE 1

Vocational Literacy Summaries and Sources

1. *Readability of Vocational Materials*
   Sticht, Tom. *Reading for Working* (1975)
      Validation of the FORECAST readability formula for technical materials
      Summary of readability of job related material.

   Sticht, T., Welty-Zapf, D. (Eds.) *Reading and Readability in the Armed Services* (1976)
      History of research and development in the Armed Services
      Ongoing research and development in the Armed Services
      Ongoing research in readability and useability in the Armed Services.
2. *Comprehensibility*
   Sticht, T. Welty-Zapf, D. (Eds.) *Reading and Readabilty in the Armed Services* (1976)
3. *Minimum Literacy Requirements of Specific Occupations*
   Moe, A., Rush ., & Storlie, R. The Literacy Requirements of 10 Occupations (1980)
      Minimum literacy requirements of account clerk, automotive mechanic, draftsman, electrician, industrial maintenance mechanic, licensed practical nurse, machine tool operator, secretary and welder studied.

   Sticht, T. *Reading for Working*
      Survey of seven occupations in the military.
4. *Job Performance and Reading Levels of Employees*
   Sticht, Tom. *Reading for Working*

   Sticht, Tom, Fox, Hauke & Zapf. *Reading in the Navy* (1976)
      Attitudes toward reading
      Determining the reading demands of Navy ratings
      Integrating job skills and reading training
      Development of Job Reading Task Tests (JRT)

   Diehl, W. & Mikulecky. The Nature of Reading at work, *Journal of Reading*, December 1980.
5. *Summaries*
   Larson, Gordon "Research on Literacy and Military Training: A Special Report to the Assistant Secretary of the Army for Manpower and Reserve Affairs." (1978)

have produced and validated a simple readability formula suitable for technical material. They have assessed reading materials in a variety of military occupations and have developed a series of pragmatic job reading task tests (JRTT) (Sticht, 1975). The same techniques were used to establish minimum literacy requirements for a census of ten civilian occupations by Moe, et al., at Purdue (1980). Sticht's work *Reading in the Navy* investigates not only the reading demands of various occupations but provides descriptive data on task frequency and the attitudes of military personnel towards reading training. Briefly, most men will take remedial reading courses if they are on company time rather than in their own time (1976).

Although much of Sticht's work is descriptive, he does provide methodologies for determining reading demands of jobs, and integrating job and reading training. Theoretically his work makes several important distinctions not made by traditional reading comprehension taxonomies. *Reading to learn* is separated from *reading to do.* As Diehl and Mikulecky (1980) point out, reading to do a job may require less skill than reading new materials in a learning situation. The reader on the job deals with familiar vocabulary in a context where extra linguistic cues, for example, the actual machinery used in the job, aid in the task of comprehending. Consequently, on-the-job reading tasks may be successfully accomplished with reading skills measuring at least two grades below those the readability of the job materials might suggest (Sacher and Duffy, 1978; Diehl and Mikulecky, 1980).

Further distinctions can be made between *reading to do* with incidental learning and *reading to assess* which involves quickly skimming material to determine its usefulness (Diehl and Mikulecky, 1980).

Usability of technical materials and comprehensability are two specific concepts associated with vocational reading. Usability refers to the ease with which workers can refer to or retrieve information from manuals. Comprehensibility goes beyond the readability of a document, covering aspects of the layout and the graphic components of the text. Most technical manuals rely heavily on the ability of the reader to extract information from figures, tables and texts. A usable comprehensibility index is provided by Muller in *Reading and Readability in the Armed Services* (1976). Sticht has researched the relative effect of extracting information from the various elements of technical manuals in his research with the Navy (1976). A detailed analysis of the research associated with both usability and comprehensibility is beyond the scope of this paper.

In summary, the research in vocational literacy can be characterized as a pragmatic attempt first to describe job related reading, to assess its difficulty, and finally to find ways of assessing the skills of the learner relative to the reading demands of the job.

Major difficulties are posed by the constantly changing demands of jobs in technical fields. What is the expected useful life of any job-related reading task test in computer electronics for example? Since the range of occupations today is so great, some effort should be directed towards clusters of critical reading skills relating to a range of technical jobs. Finally, the bottom line for much of this research is productivity. Will the recipients of intensive job-related reading programs actually perform better on the job?

PART III: SUMMARY AND IMPLICATIONS

The current trend in instruction has been away from teaching a traditional academic curriculum based on the skills and content taught in grade school and towards teaching functional life skills. Current definitions of literacy also stress functional life skills. Adequate literacy in Adult Performance Level terms is a composite of income, education and job status (Kirsch and Guthrie, 1977-78).

This trend is worrying both in philosophical and practical terms. The humanities do have a place in the adult literacy classroom. People should not be denied a liberal education because of age or income. Since it is the poor who are most subject to illiteracy, it is they who get the life skills "how to read a bus schedule" option, while the rest get liberal arts.

The same arguments are used against functional life skills programs that were applied to vocational education (Lazerson and Grubb, 1978). "Vocational" became synonymous with second class status—a place for those who couldn't make the academic track.

On a practical level, the scope of a functional life skills curriculum such as the APL is overwhelming. There is a real danger that overworked teachers will select out the least complex skills, those that are almost trivial. They may teach what they perceive to be those skills most useful to the student rather than those skills with maximum generalizability. Students may need guidance in selecting not only those skills that coincide with their interests, but those that are most useful, not just immediately applicable.

In the area of teacher training the overwhelming issue seems to be the marginal part-time status of teachers and the lack of any coherent teachers' preparation such as might be embodied in licensure standards. Adult literacy is still considered a marginal field taught by moonlighters who have had little preparation for their jobs. Since they are paid by the hour there is little incentive to participate in inservice training. Hunter and Harmon (1979) attribute the high dropout rate in literacy programs, in part, to poor teaching.

Finally, although the demographic data is unclear as to the extent of the literacy problem—partly because we have no general agreement as to what illiteracy is—there is little doubt as to who the illiterate are. We have a mass of descriptive data about them. Perhaps the time has come to concentrate on smaller, well controlled studies, even in-depth case studies. For example, it might be productive to study the learning disabled adults who have learned to read—the compensatory strategies they use might help other adults learn. At the University of Minnesota we are beginning a series of case studies of illiterate families. We are finding that, perhaps due to lack of both literacy and schooling, illiterate adults group and classify words difficiently when compared to literate peers. We are now looking at the ability of their children. In the vocational area particularly, we need to test the effectiveness of teaching techniques in controlled experimental situations. The military is ideally suited to this.

Finally we need to establish the legitimacy of the adult literacy field. Research is difficult, subjects are hard to locate and often unwilling. However, the researcher is provided with vast opportunities—so little has been done and there is so much to do.

REFERENCES

American Bar Association Commission on Correctional Facilities and Services. *Offender Literacy Project.* Washington, D.C.: ABA Clearinghouse, 1971.

Anderson, D. and Neimi, J. Adult education and the disadvantaged adult. *Syracuse University Publications in Continuing Education,* Number 22.

Bauer, B. *A comparison of the performance of adult students in competency based and traditional academic curricula.* Unpublished Doctoral Dissertation, University of Minnesota, 1980.

Beach, R., Bernazza-Hasse, A. and Robinson, R. Teaching the aged reader. *Educational Gerontology,* 1979, *4,* 253-259.

Buswell, C. *Reading and the Aged.* Wilson Library Bulletin, *45,* January 1971.

Cadwell, R. *The acquisition of language skills using computer based instruction.* Paper presented at the Symposium on Cognition and Written Language, Pittsburgh, Pennsylvania, March 1978.

Dauzat, S. *Louisiana Adult Performance Level Pilot Study: A Comparative Analysis of APL Competency-Based Instructional Programs.* Ruston: Louisiana Tech University, 1977.

Darkenwold, G. Some effects of the obvious variable. Teacher's race and holding power with Black adult students. *Sociology of Education,* 1975, *48,* 420-431.

Diehl, W. and Mikulecky, L. The nature of reading at work. *Journal of Reading,* 1980, *24,* 221-227.

Even, M. J. *Adult Basic Education Teacher Survey.* Lincoln: University of Nebraska, 1975.

Eyster, G. *Final Report: Appalachian Adult Education Center.* Morehead, Kentucky: Morehead University, 1974.

Flaherty, J. *A Study of Functional Education Needs of New Jersey Adult Basic Education Studies.* Rutgers: Center for Adult Development, Rutgers University, July 1977.

Gallup Poll. New York: American Institute of Public Opinion, 1969-1976.

Glynn, S. and Muth, D. Text learning capabilities of older adults. *Educational Gerontology*, 1979, *4*, 253-269.

Greenberg, B. and Dervin B. *Communication Among the Urban Poor: Sample Characteristics.* Report Number 3. Lansing: Michigan State University, Department of Communications, 1968.

Griffith, W. and Cervero, R. *The Adult Performance Level Program: A Serious and Deliberate Examination.* Address presented at the Adult Education Act/National Advisory Council on Adult Education Convention, New York, November 1970.

Harris Poll. *The 1971 Reading Difficulty Index.* New York, 1971.

Hunter-Carman, St. John, and Harmon, D. *Illiteracy in the United States.* McGraw Hill, New York, 1979.

Hutchinson, L. F. The relationship between expressed interests and reading achievement in functionally illiterate adults. *Reading Improvement*, 1978, *14*, 203-207.

Jencks, C. *Inequality.* New York: Basic Books, 1972.

Kirsch, I. and Guthrie, J. The concept and measurement of functional literacy. *Reading Research Quarterly*, 1977-78, *XIII*, *4*, 485-507.

Kling, M. and Tivenan, B. Reading activities of black adult education students. *Adult Education*, 1978, *28*, 156-164.

Kohlberg, L. Stage and sequence: The cognitive developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of Socialization Theory and Research.* Chicago: Rand McNally, 1969.

Kozol, J. M. A new look at the literacy campaign in Cuba. *Harvard Educational Review*, 1978, *48*, 341-378.

Lazerson, A. and Grubb, N. Rally round the workplace: The fallacies in career education. *Harvard Educational Review*, 1975, *45*, *4*, 451-474.

Manning, D. T. Everyday materials improve adults' reading. *Journal of Reading*, 1978, *21*, 721-724.

Mezirow, J., Darkenwold, G. and Knox, A. *Last Gamble on Education.* Washington, D.C.: Adult Education Association, 1975.

Mikulecky, L., Shanklin, W. and Caverly, D. Adult reading habits, attitudes and motivations: A cross sectional study. *Monographs in Teaching and Learning*, 1979, Bloomington: Indiana University.

Mocker, D., et al. *Adult Basic Education Teacher Competency Inventory.* Kansas City: University of Missouri, Center for Resource Development in Adult Education, 1974.

Moe, A., Rush, T., and Storlie, R. *The Literacy Requirements of: account clerk, welder, automotive mechanic, draftsman, electrician, heating and air conditioning mechanic, industrial maintenance mechanic, licensed practical nurse, machine tool operator, secretary.* Department of Education, Purdue University, June 1980.

Murphy, R. T. Assessment of adult reading competence. In D. M. Nielsen and H. F. Hjelm (Eds.), *Reading and Career Education: Perspectives in Reading.* (19). Newark, Delaware: IRA, 1975.

National Advisory Council on Adult Education. *A Target Population in Adult Education.* Washington, D.C.: Author, 1974.

National Assessment of Educational Progress. *NAEP Newsletter. X*, February 1977.

Northcutt, N. *The Texas Adult Performance Study.* Final Report. Austin: University of Texas, 1975.

Osso, N. A. *Adult Basic and Secondary Level Program Statistics.* Washington, D.C.: National Center for Educational Statistics, 1975.

Park. R. Women in adult basic education. *Lifelong Learning: The Adult Years.* December 1977, 12-13.

Park, R. *An Evaluation of the Use of the PLATO Basic Skills Learning System in Minnesota.* Bloomington, Minnesota: Control Data Corporation for Bloomington, 1978.

Resnick, D. and Resnick, L. B. The nature of literacy: An historical exploration. *Harvard Educational Review*, 1977, *47*, *3*, 370-385.

Robinson, R. D. Literacy instruction—an opinion survey of persons in adult education. *Adult Leadership*, 1974, *23*, 167-169.

Rossman, M. *A Model to Recruit Functionally Illiterate Adults into Adult Basic Education Programs in Massachusetts.* Amherst: University of Massachusetts, 1971.

Sacher, J. and Duffy, T. *Reading skill and military effectiveness.* Paper presented at the American Educational Research Association Conference, Toronto, Canada, March 1978.

Seaman, D. *Preventing Dropouts in ABE.* Tallahassee: Florida State University, Adult Education Information Processing Center, 1971.

Sharon, A. T. What do adults read? *Reading Research Quarterly*, 1973-1974, *IX*, 148-169.

Spady, W. Competency based education: A bandwagon approach. *Educational Researcher*, 1977, *6*, 9-14.

Spear, G. *Adult Basic Education National Teacher Training Study Part III Survey of Needs*. Kansas City: University of Missouri, Center for Research Development in Adult Education, 1972.

Sticht, T. *Reading in the Navy*. Alexandria, Virginia: Human Resources Research Organization, September 1976.

Sticht, T. *Reading for Working*. Alexandria, Virginia: Human Resources Research Organization, 1975.

Ulmer, C. *Teaching the Disadvantaged Adult*, National Association for Public and Continuing Adult Education. Washington, D.C., 1969.

U.S. General Accounting Office. *A Need to Address Illiteracy Problems in the Military Services*. (U.S. FAO FPCD 77 1B) Washington, D.C.: Author, March, 1977.

Wisher, R. *CAI and Naval Educational Strategies in Basic Skills*. Paper presented at the 25th Convention of the I.R.A., St. Louis, Missouri, May 1980.

Wool, H. and Flyer, E. S. *Project 100,000: Characteristics and performance of New Standards for Men*. Washington, D.C.: U. S. Department of Defense, Office Secretary of Defense, Assistant for Manpower and Reserve Affairs, 1968.

Yankelovich, S. and White. *Exerpt of Selected Findings From the Consumer Study on Reading and Book Purchasing*. The Book Industry Study Group, 1978.

Zeigler, W., Healy, G. and Elksworth, J. Civic Literacy in Klevens, C. (Ed.) *Materials and Methods in Continuing Education*, Klevens Publications, Inc. New York, 1976, 130-150.

PROGRAM

The 1980 National Reading Conference[1]

3.1 Anne Forester; Camosun College. **Preparing a Reading and Study Skills Package for Distance Education**

3.2 Ann Coil; California State University-Fullerton. **The Perry Scheme: Applications to College Reading.**

4.1 Bess Altwerger; Arizona State University. **Processing of Metaphor.**

4.2 Claudia Dybdahl, University of Arizona, **Linguistic Awareness and Literacy.**

4.3 Barbara Flores, University of Arizona, **Psycholinguistic Principles in Literacy and Biliteracy.**

4.4 Yetta Goodman, University of Arizona. **A Model of Beginning Reading.**

5.1 Jane Hansen, University of New Hampshire. **An Intervention Study; Improving the Inferential Comprehension of Fourth Grade Good and Poor Readers.**

5.2 Patricia Herman, Toledo Public Schools, Peter Dewitz & John Stammer, University of Toledo. **The Development of Syntactical Chunking in Intermediate Grade Non-Fluent Readers Using the Method of Repeated Readings.**

5.3 Linda McHugh & Ray Buss, University of Wisconsin-Madison. **Diagnostic and Prescriptive Decisions made by Specialists and Teachers: Quadratic Assignment as a Reliability Measure.**

5.4 Dixie Spiegel & Jill Whaley, University of North Carolina-Chapel Hill. **Elevating Comprehension Skills by Sensitizing Students to Structural Aspects of Prose.**

8.1 Linda Johnson, University of Wisconsin. **The Effects of Stylistic Factors on Textbook Readability.**

8.2 Terry Salinger, New Mexico State University & Emily Payne, Austin Commercial College. **A Developmental Approach to Curriculum: College Reading and Study Skills.**

8.3 Dennis Kear, Michael McKenna & John Miller, Wichita State University. **An Examination of the Efficiency of Four Reading/Study Techniques.**

8.4 Tina Jacobowitz & Michael Maslanek, Montclair State College. **The Effects of Previewing Expository Text on Recall & Recognition.**

10.1 Peter Dewitz & John Stammer, University of Toledo. **The Development of Linguistic Awareness in Young Children from Label Reading to Word Recognition.**

10.2 Ruth Cook, Southern Illinois University. **A Study of the Strengths and Weaknesses in Readiness Skills Typical of Children Entering Kindergarten.**

10.3 Judith Greenwald, City University of New York. **Children's Concept of Temporal Sequence and Their Comprehension of Inverse Temporal and Causal Sentences.**

10.4 Sarah Rogers, Corpus Christie State University. **A Description of Four Stages of Oral Reading Behavior Exhibited During the First Six Weeks of Learning to Read.**

13.1 Charles Lindamood, Patricia Lindamood, & Phyllis Lindamood, Lindamood Center. **The Auditory Conceptual Factor in Literacy Development.**

13.2 Miriam Pepper-Sanello, C.D.S. No. 19, Brooklyn, NY. **The Effect of Auditory Distraction at Varying Intensities on Listening and Reading Ability.**

14.1 Charles Kinzer, University of California-Berkeley. **Questions from an Interdisciplinary Perspective: Some Answers are Already out There.**

14.2 Karen Carroll, University of California-Berkeley. **Questions from a Temporal Perspective: Adding a Missing Dimension to Reading Theory.**

14.3 Donald Leu, University of California-Berkeley. **Questions from a Metatheoretical Perspective: The Interdependence of Possible Solutions.**

15.1 Susan Goldman & Connie Varnhagen, University of California-Santa Barbara. **The Development of Listening and Reading Comprehension: The Role of Story Ending Information.**

15.2 Joyce Hood, University of Iowa. **Procedures for Scoring Retellings, Their Reliability and Sensitivity to Effects of Prose Type and Growth in Ability.**

15.3 Andee Rubin, Bolt, Beranek & Newman. **The Story Maker in the Classroom: A Follow-up.**

15.4 William Teale, University of California-San Diego. **How Preschoolers Interact with Written Communication: First Year Report.**

18.1 David Caverly, Indiana University. **An Application of Two Strategies for Improving Reading Comprehension and Retention.**

18.2 Janet Jackson, Indiana University. **Assessing Student Learning and Retention Strategies Using Retellings of Text Passages.**

---

[1] As presented, the program includes all of those papers listed in the official conference program.

18.3 Diane Menendez, University of Cincinnati. **Reading and Rereading, Writing and Re-writing: Success of Revision Strategies Used by Basic Students.**

19.1 Darrell Morris, National College of Education. **Concept of Word and Sound Segmentation Ability.**

19.2 Jerry Zutell, Ohio State University. **Cognitive Development, Metalinguistic Ability, and Invented Spellings: Comparisons and Correlations.**

19.3 Tobie Sanders, Ohio State University. **Children's Concept of Word and Language of Instruction.**

19.4 Shane Templeton, Emory University, & Elizabeth Sulzby, Northwestern University. **Beyond the Psycholinguistic Vise of Competence/Performance Theory: Why Study Metalinguistic Awareness?**

19.5 Elizabeth Sulzby, Northwestern University, & Shane Templeton, Emory University. **The Development and Validation of Metalinguistic Tasks Assessing Children's Concepts of "Word-ness".**

20.1 James Juola, Nicklas Ward, Hsuan-Chih Chen, & Cindy Ikenaga, University of Kansas. **Reading Serially-presented Text.**

20.2 Ron Leslie, New York University. **The Utilization of an Augmented Transitional Network to Investigate Early Reading Processes.**

23 Roger Shank, Yale University. **Reading and Understanding: A Perspective From Artificial Intelligence**.

25.1 Dale Johnson, Susan Pittelman, Judy Schwenker, & Linda Shriberg, University of Wisconsin-Madison. **A Focus on Mastery: A New Approach to the Assessment of Word Identification Skills.**

25.2 Carmen Andrew, University of Iowa. **The Effect of Vocabulary Difficulty and Contextual Predictability on the Reading Efficiency of Beginning Readers.**

25.3 Connie Juel, University of Texas-Austin, & Robert Solso, University of Idaho. **The Role of Orthographic Redundancy, Versatility, and Spelling-Sound Correspondence Used in World Identification.**

25.4 Leon Williamson, New Mexico State University. **Facets of English Spelling: Phonics, Semantics, Etymology, Preference.**

26 Harry Singer, University of California—Riverside. **Hypotheses in Search of Classroom Validation**.

27.1 David Elkind, Tufts University, & Dawn Mayhew, University of Montana. **Figurative Play and Beginning Reading.**

27.2 Dawn Mayhew, University of Montana, & David Elkind, Tufts University. **Equivalence Decoding: A Reading Readiness Consideration.**

28.1 Edward Fry, Rutgers University. **A Partial Reading Model Utilizing Language Unit Size by Frequency.**

28.2 Richard Hodges, University of Pudget Sound. **Phoneme-Grapheme Frequency Counts.**

28.3 Elizabeth Sakiey, Glassboro State University. **Syllable and Word Frequency Counts.**

28.4 Josephine Goldsmith, Rutgers University. **Frequency of Larger than Word Units.**

29.1 Richard Robinson, University of Missouri, & Ann Haase, University of Arizona. **Diagnostic and Interest Reading Inventories of Middle Aged Reading Specialists and Sophisticated Readers.**

29.2 Paulette O'Malley & Ann Haase, University of Arizona. **Diagnostic and Interest Reading Inventories of GED and Community College Students.**

29.3 Molly Wilson, Idaho State University. **Reading Interests of the Elderly.**

29.4 Alice Wescott, Phyllis Berlowe, & Ann Haase, University of Arizona. **Cloze Error Profiles of Spanish Speaking American Adults.**

29.5 Stephanie Brown, University of Arizona. **An Examination of the Subject Matter Preference of Adults Concerning Reading Material for Children and Adolescents.**

29.6 Frederick Casler, Brock University. **Adult Reading of Ancient Languages.**

30.1 John Readence, University of Georgia. **Reviewing Applied Research in Comprehension Instruction.**

30.2 David Moore, University of Georgia. **Some Suggested Approaches and Procedures for Researching Comprehension Instruction.**

30.3 Richard Barron, Oakland University. **An Example of Systematic Research in Content Area Comprehension Instruction.**

31.1 Linda Gambrell & Patricia Koskinen, University of Maryland, & Jack Cole, Prince George's County Public Schools. **Metacognitive Knowledge about Mental Imagery: A Comparison of Above and Below Average Readers.**

31.2 Rebecca Palmer, University of Minnesota. **How Good Studiers Handle Written Discourse.**

31.3 Gail Huffman, Michael Green, & Barbara Edwards, University of North Carolina-Charlotte. **Development Stages of Metalinguistic Awareness Related to Reading.**

31.4 Michael Green, Gail Huffman, & Barbara Edwards, University of North Carolina-Charlotte. **Development Stages of Metalinguistic Awareness Related to Oral Communication.**

32.1 Janet Kendall, Simon Fraser University. **Literature Review: Research Problems.**

32.2 Jana Mason, University of Illinois-CSR. **Instructional Questions re. Generic Vocabulary Skills.**

32.3 Jaap Tuinman, Simon Foster University. **A Research Plan to Investigate the Development of Polysemy.**

33.1 Ellen Ryan, University of Notre Dame. **Methodological Problems in Research Strategy Defficiencies among Poor Readers.**

33.2 Margaret MacLean, McGill University. **Comprehension of Incomplete Messages.**

33.3 Carol Latour, Lakeshore School Board, Montreal. **Progressive Cloze: Qualitative Insights into Narrative and Expository Text Processing.**

33.4 Peter Winograd, University of Illinois-CSR. **Examination of Summaries Produced by Good and Poor Readers.**

33.5 Mark Aulls, McGill University. **Developmental Considerations of Reading Research on Good and Poor Readers.**

35.1 Martha Haggard, Northern Illinois University, & Robert Ruddell, University of California-Berkeley. **The Inservice Instructional Cycle (IIC): A Paradigm for Conducting Inservice Programs.**

35.2 Mary Geissal, Northeastern Illinois University, June Knafle, & Susanna Pflaum, University of Illinois-Chicago. **The Affective Responses of Teachers and Pre-Teachers to Spelling Variations.**

35.3 Lee Mountain, University of Houston. **A Graduate Program in Reading that Helps Students Publish.**

36.1 Judith Langer, New York University, & Mark Nicolich, Millstone, N. J. **Effect of Altered Prior Knowledge Level on Passage Recall.**

36.2 Gregory McKenzie, Allan Neilsen, & Carl Braun, University of Calgary. **Effect of Linguistic Connectives and Prior Knowledge on the Understanding of Causal Relationship by Good and Poor Readers.**

36.3 Peter Johnston & David Pearson, University of Illinois—CSR. **Explicitness of Connectives and Content Familiarity as Determinants of Reading Comprehension.**

36.4 Jennifer Stevenson, University of Wisconsin-Madison. **The Effects of Explicit-Activated and Implicit-Activated Antecedents on Average Third and Eighth Grade Readers' Resolution of Anaphora.**

Rosemarie Park, University of Minnesota. **A Critical Review of Developments in Adult Literacy.**

39.1 Philip Lanasa, University of Houston-Clear Lake City. **Comparison of Functional-Survival Core Vocabularies for the Adolescent Slow Learner: Research Basis vs. Empirical Basis.**

39.2 Sandra Stotsky, Curry College. **Toward Reassessment of the Principles Underlying Choice of Vocabulary and the Teaching of Word Analysis in Reading Instructional Material.**

39.3 Michael Graves & Wayne Slater, University of Minnesota. **Some Measurements of the Depth of Children's Word Knowledge.**

40.1 Diane Lapp, San Diego State University. **Utilizing Questions as a Comprehension Stimulant.**

40.2 James Flood, Boston University. **Metalinguistic Awareness: The Teachable Part.**

40.3 Betty Holmes & Nancy Roser, University of Texas-Austin. **A Comparison of Techniques Used to Assess Reader's Prior Knowledge.**

40.4 Robert Calfee & Robert Curley, Stanford University. **Comprehending and Knowing: Formal and Informal Discourse Structures.**

40.5 Priscilla Drum & Carol Dixon, University of California-Santa Barbara. **Observation of Student-Teacher Interaction in Reading Comprehension Instruction.**

40.6 David Pearson, University of Illinois—CSR. **A Review of Instructional Research in Reading Comprehension.**

40.7 Carol Dixon & Priscilla Drum, University of California-Santa Barbara. **Application of Discourse Analysis Research to Diagnosis.**

40.8 Robert Tierney, University of Illinois—CSR., Diane Schallert, University of Texas-Austin, Jill

LaZansky, University of Illinois—CSR. **Aspects of Incoherence in Biology and Social Studies Text.**

41.1 Alice Boljonis, Syracuse University, & Daniel Kaye, Yale University. **Differences in 4th Grade Readers' Recalls of Written Prose as an Effect of the Type and Level of Structure Present in the Text.**

41.2 James Baumann, University of Wisconsin-Madison. **Effects of Height of Information in the Content Structure and Measurement Mode on Third and Sixth Graders' Comprehension of Expository Discourse.**

41.3 Barbara Rennie, Allan Neilsen, & Carl Braun, University of Calgary. **The Effect of Typographical Cueing on Memory for Superordinate Relationships in Written Discourse.**

41.4 Rosalind Horowitz, Gene Piché, & Jay Samuels, University of Minnesota. **A Study of the Effect of Contrasted Rhetorical Predicates on the Recall of Expository Discourse.**

42.1 Kenneth Goodman, University of Arizona. **The Right View.**

42.2 Philip Gough, University of Texas-Austin. **The Wrong View.**

43.1 Robert Carey, Indiana University. **Naturalistic Inquiry: The Non-Dominant Paradigm.**

43.2 Jerome Harste, Indiana University. **Classroom Constraints on Comprehension Instruction.**

43.3 Karen Feathers, Indiana University. **Comprehension as Extended Text.**

43.4 Margie Siegel, Indiana University. **Discourse Processing from the Perspective of Speech Act Theory.**

44.1 Anne Wolf, University of Wisconsin-Eau Claire, & Jane Greenewald, University of Wisconsin-La Crosse. **Frequency of Reading in Secondary Content Areas: A Follow-up Observation Study.**

44.2 Ken Dulin & Mary Sue Dillingofski, University of Wisconsin-Madison. **Five Secondary-Level Reading Attitude Measures compared with Various Teacher-Selected Criterion Measures of Reading Ability.**

44.3 Shirley Merlin, James Madison University, & Nadja Sefeik-Earl, Petersburg High School, WV. **The Relationship between Reading Frustrations and Behavior Characteristics of Good and Poor Readers.**

44.4 John Readence & David Moore, University of Georgia. **An Investigation into the Effects of a Student-Directed Categorization Strategy on Concept Development.**

46.1 John Miller, Wichita State University. **Differences Between Low Income Schools with High Reading Achievement and High Income Schools with Low Reading Achievement.**

46.2 Lea McGee, Virginia Polytechnic Institute. **Good and Poor Readers' Ability to Distinguish Among and Attend to Levels of Importance in Expository Text.**

46.3 Helen Gabriel, Carl Braun & Allan Neilson, University of Calgary. **The Effects of an Author's Textual Scheme on the Comprehension of Good and Poor Readers.**

46.4 Susanna Pflaum & Ernest Pascarella, University of Illinois-Chicago. **The Contribution of Prior Knowledge to Disabled and Normal Readers' Comprehension.**

47 Ann Brown, University of Illinois—CSR. **Writing and Reading and Metacognition.**

48.1 Thomas Estes, University of Virginia. **Properties of Text Predicting Comprehension.**

48.2 Mary Wetmore, University of Virginia. **Does Cueing Improve Recall?**

48.3 James Deese, University of Virginia. **Form of the Text and Comprehension.**

48.4 Wyane Shebilske, University of Virginia. **Study Strategies and Comprehension.**

48.5 John Rotondo, University of Virginia. **Mapping the Perceived Structure of Text: Methodological Prospects.**

49.1 Beth Evans, Laramie County School District No. 1, Wy. **Developmental Trends in a Retention of Proposition Level.**

49.2 Allan Neilsen, University of Calgary. **Effects of Micro-and Macro-Structure on Text Processing.**

49.3 Barbara Taylor, University of Minnesota. **The Effects of Instruction in Following Text Structure on Children's Memory for Expository Discourse.**

49.4 Elizabeth Sulzby & Pamela Adelman, Northwestern University. **The Recall of Text Structure as Affected by Inferences Made During Reading.**

49.5 Victor Froese, University of Manitoba. **Exploratory Studies in the Practical Application of Discourse Analysis: Informal Inventories and ESL Learning.**

50.1 Edys Quellmalz, University of California-Los Angeles. **Assessing Writing Proficiency.**

50.2 Eva Baker, University of California-Los Angeles. **Assessing Reading Proficiency.**

50.3 Frank Capell, University of California-Los Angeles. **Issues in Assessing the English Writing Performance of Limited English Speaking Students.**

50.4 William Cooley, University of Pittsburgh. **Assessing Reading Proficiency of the Learning Disabled.**

51.1 William Smith, Indiana University. **The Use of Story Grammars in the Development of a Model of Discourse Processing.**

51.2 Nora Hoover, Virginia Polytechnic Institute. **The Acquisition of a Story Schema.**

51.3 Nancy Stein, Northwestern University. **How Children Tell Stories: Theoretical and Empirical Concerns.**

52.1 Eugene Cramer, University of Illinois-Chicago, & Cheryl Hendress, Bureau County ESR, Pegasus, IL. **Analogical Reasoning, Mental Imagery, and Reading Ability of Middle Grade Students: A Correlational Study.**

52.2 Deborah Cureton, University of Wisconsin-Madison. **Differential Processing of Explicit and Implicit Information by Skilled and Difference Fourth Grade Readers.**

52.3 David Haynes, University of Georgia. **Latent Transfer Effects of Analogy Upon Prose Learning.**

52.4 Elizabeth Taylor, New York State Education Development. **Confirming Story Grammar Structures.**

55.1 Christine Lucas, McGill University. **Passage from Reading Patterns to Writing Patterns Supported by Linguistic Connectives.**

55.2 Frank Torres, California State Polytechnic University, & Leon Williamson, New Mexico State University. **College Freshmen's Sensitivity to Psycholinguistic Cues when Reading.**

55.3 James Pichert, Vanderbilt University. **Adult Sensitivity to Relative Text Importance.**

55.4 Norma Inabinette & Ruth May, California State University-Fullerton. **Adult Cognition: Piaget Revised.**

57.1 Nancy Rogers-Zegarra, University of California-Riverside. **Anglo and Chicano Literal and Scriptal Comprehension of Ethnic Stories.**

57.2 Thomas Bean & Ashley Bishop, California State University-Fullerton. **Selected Prose Characteristics of ESL Materials and Their Effect on Comprehension.**

57.3 Mark Condon, University of Louisville. **Secondary Reading Inventory: Source of Questions.**

57.4 Ronald Carver, University of Missouri-Kansas City. **Just How Flexible are Readers?**

58 Gerald Duffy, Michigan State University. **Teacher Effectiveness Research.**

59.1 Connie Bridge & Susan Thomas, University of Kentucky. **The Effects of Textual Cohesion on Cloze Comprehension Scores.**

59.2 Ernest Goetz, Texas A & M University, Karen Dixon, Robert Kantor, & Tsivia Cohen, University of Illinois—CSR. **Patterns of Context Utilization by Good and Poor Readers.**

59.3 Michael McKenna, Wichita State University. **Practice Effect of Cloze Testing.**

59.4 Earl Rankin, University of Kentucky & Ann Haase, University of Arizona. **The Effects of Three Scoring Procedures on the Reliability & Validity of Cloze Reading Comprehension Tests for Mexican-Americans.**

62.1 Jill Whaley, University of North Carolina-Chapel Hill. **Reader's Reactions to Temporal Disruption in Stories: Another Revelation of Expectations for Story Structure?**

62.2 Kathleen Gormley & Mary Marr, State University of New York-Albany. **Comprehension: A Look at Reading Ability and Experience with Text.**

62.3 Candace Bos, University of Arizona. **Remembering Information from Text: A Comparison of the Text Structure of EMR and Average Student's Recalls.**

62.4 Sandra Berkowitz & Barbara Taylor, University of Minnesota. **The Effects of Text Type and Familiarity on Readers' Recall Mode.**

64.1 Nancy Snouffer & Linda Thistlethwaite, Western Illinois University. **The Effect of Graphic Pre and Post Organizers upon the Comprehension of College Students Reading History Material.**

64.2 Charles Clark, Western Illinois University, & Thomas Bean, California State University-Fullerton. **Objectifying the Construction of Advance Organizers: Possible but not Necessarily Worthwhile.**

64.3 Charlene Swanson, California State College-San Bernardino. **Effects of Readability and Organizational Pattern on Ninth Graders' Written Recall.**

64.4 Michael Johnston, Tucson Unified School Dist. No. 1, Wilbur Ames, & Frank Spencer, University of Arizona. **The Effect of Manipulating the Advance Organizer and Other Prereading Strategies on the Comprehension of Abstract Text.**

65.1 Robert Schwartz, Oakland University. **Coordination and Monitoring of Information in the Word Recognition Decisions of Good and Poor Readers.**

65.2 Peter Winograd & Peter Johnston, University of Illinois—CSR. **Comprehension: Monitoring and the Error Detection Paradigm.**

65.3 Taffy Raphael & Robert Tierney, University of Illinois—CSR. **The Effect of Topic Familiarity and Discourse Force on Detection of Inconsistent Information.**

66.1 Brenda Wexler, Carol Pierson & Patricia Anders, University of Arizona. **Cleaning Up the Reading and Writing Act: Developing Functional Reading and Writing Competencies in Middle Management.**

66.2 Mary Marr & Kathleen Gormley, State University of New York-Albany. **Factors that Influence Undergraduates' Reading Comprehension.**

66.3 Emily Mager, Instructional Science and Development, Inc. **Document Design and Visual Representations of Text as Aids in Comprehension.**

66.4 Mary Heller, University of Toledo. **The Relationship between Reading Comprehension and Written Expression.**

67.1 Michael Kamil, University of Illinois-Chicago. **Imagery and Frequency in Word Recognition Models.**

67.2 Glen Powell, Kearney State College. **A Meta-Analysis of the Effects of Imposed and Induced Imagery upon Word Recall.**

67.3 Tess Kosoff, Syracuse University. **The Effects of Three Types of Introductions on Children's Processing of Text.**

67.4 Karen Wixson, Syracuse University. **The Effects of Post Reading Questions on Children's Discourse Comprehension and Knowledge Acquisition.**

68 Richard Venezky, University of Delaware. **Letter Sound Regularity and Othographic Structure.**

69.1 Cynthia Moore, Reading Consultant, Region IV, Houston. **Results of Need Assessment, Pre and Post Criterion Referenced Test.**

69.2 Betty Criscoe, University of Houston-Clear Lake City. **The 8 Teacher Training Sessions.**

69.3 Jake Parker, Reading Consultant, Region IV, Houston. **Six Training Films for Physical Science Teachers.**

70.1 Karen Dixon, Andrew Ortony & David Pearson, University of Illinois—CSR. **Some Reflections on the Use of Figurative Language in Children's Books.**

70.2 Stella Vosniadou, University of Illinois—CSR. **The Conception of Literal and Metaphorical Similarity in the Pre-School Child.**

70.3 Marshá Taylor, University of Illinois—CSR. **Figurative Language in Black Language Patterns: Educational and Pedagogical Implications.**

70.4 Stella Vosniadou, Ralph Reynolds, & Paul Wilson, University of Illinois—CSR. **The Comprehension of Metaphor in the Pre-School Child.**

71.1 Martha Haggard, Northern Illinois University. **Group Mapping Activity-Secondary Level.**

71.2 Bonnie Greenslade, Northern Illinois University. **Group Mapping Activity-Analysis of Students' Questions and Responses.**

71.3 DeWayne Triplett, Northern Illinois University. **Group Mapping Activity-Analysis of Individual Mapping Styles for Four Types of Literature Passages.**

72.1 Jerome Niles, Virginia Polytechnic Institute. **Effects of Teacher Feedback on Beginning Reader's Oral Reading Behavior.**

72.2 James Hoffman, University of Texas-Austin. **A Comparison of Pre-Service and In-Service Teacher Feedback Strategies During a Dyadic Oral Reading Interaction.**

72.3 Christopher Baker, University of Texas-Austin. **The Nature and Distribution of Sustaining Teacher Feedback During Oral Reading.**

72.4 Sharon O'Neal, University of Texas-Austin. **Teacher Feedback During Oral Reading and Its Relation to the Subsequent Behavior of the Reader.**

73.1 Richard Barron, Oakland University. **Graphic Post-Organizers, Comprehension and Transfer in a Complex, Long-Term Learning Task.**

73.2 Joseph Vaughan, East Texas State University. **Structured Overviews and Schema Building.**

73.3 Gary Long, Rochester Institute of Technology. **Dissecting Prose Passages Through Semantic Mapping: Procedures for Training.**

73.4 Bonnie Armbruster, University of Illinois—CSR, & Diane Schallert, University of Texas-Austin. **Understanding Text Through Mapping.**

74.1 William Diehl, University of Georgia. **Functional Literacy: Demands and Strategies in a Cross Section of Occupations.**

74.2 Larry Mikulecky, Indiana University. **Job Literacy: The Relationship Between What Students Learn and What Workers Do.**

74.3 Gordon Larson, Rutgers University. **Reading to Do and Reading to Learn: A Literature Review and Model Presentation.**